Journal of Educational Research

...ral Office, Urbana, Illinois Editor-in-chief B. R. Buckingham, University of Illinois

...LUME II JUNE, 1920 NUMBER 1

GETTING TEACHERS TO FEEL THE NEED FOR STANDARDIZED TESTS[1]

SAMUEL S. BROOKS
Superintendent of Schools, Silver Lake, New Hampshire

Having concluded that standardized tests would greatly help ... under the conditions set forth in my first article, I next at...ked the problem of securing the whole-hearted cooperation of my ...chers in their use. The psychology of interest teaches us that ...person's best efforts are called forth only when he feels a real ...ed. So I set myself to arouse in my teachers an enthusiastic ...erest by bringing them to feel a real need for definite standards ...accomplishment in school work. In order to make them feel ...s need deeply, something more than merely telling them of the ...ts and their uses was necessary. They must be made to realize ...e inadequacy of ordinary methods of measuring results of ...ching.

Accordingly, I called a series of teachers' meetings on four con...utive Saturday afternoons. In a district more than twenty ...es square it is not easy to get all the teachers together at one ...ce. I succeeded, however, in persuading two automobile own... in each town to take the teachers to these four meetings at a ...e that would little more than pay for the gasoline consumed. ...ether or not this scheme could be worked in other districts ...ld depend, of course, on the character of the automobile ...ners and the success of the superintendent in convincing them ...t they ought to be willing to do it in an emergency and for the ...fare of the schools.

At the first meeting I outlined briefly what I wanted to do and ...y, and then suggested that we try some practical experiments.

[1] This is the second article by Superintendent Brooks on the general topic "Putting ...dardized Tests to Practical Use in Rural Schools."

The First Meeting

The first experiment was intended to prove the inaccuracy teachers' marks in general. A few days before I had selected paper turned in by a sixth-grade pupil in a regular arithme test given by one of the teachers. This test contained the us ten problems and was supposed to measure knowledge of perce age. A hektographed copy of the pupil's paper was given out each of the 24 teachers. Then, without warning them of purpose or of what the results were apt to be, I asked them correct the paper and mark it on a percentage basis. When had finished, the papers were collected and I at once tabulat the results on the blackboard with the result shown in Table I.

TABLE I. DISTRIBUTION OF RATINGS OF A SIXTH-GRADE ARITHMETIC PAPER

Ratings	Frequency
90–94	3
85–89	8
80–84	6
75–79	4
70–74	2
65–69	1

The effect was striking. Nearly everybody had caught th idea and the expressions on the various faces showed me that the was no need of rubbing it in by pointing out the moral. Arous interest was evident as I passed the papers back with the su gestion that we analyze the methods of marking to discover t factors which produced such evident lack of agreement. W finally agreed on the following points:

1. Most had marked the pupil's response to a problem entirely wrong the answer was wrong without regard to correctness of principle.
2. Some gave half credit if the principle was right and the answer wro
3. The majority marked answers entirely wrong if a decimal point w omitted or misplaced while some gave varying degrees of credit if the dig of the answer were correct and the decimal point misplaced or omitted.
4. A few had assigned weights to the various problems on the basis of th own judgments with far from uniform results. This factor actually seem to have been the one which had produced the most extreme variations in t marking.

Then someone suggested that we try it again and see if we couldn't do better. Several others seconded the idea, so I passed around copies of an eighth-grade pupil's history paper. Never did teachers work more seriously than did those teachers for the next fifteen or twenty minutes. The papers were then collected and the marks tabulated as before with these results:

TABLE II. DISTRIBUTION OF RATINGS OF AN EIGHTH-GRADE HISTORY PAPER

Ratings	Frequency
90–94	1
85–89	3
80–84	10
75–79	5
70–74	2
65–69	3

Then followed a lively discussion, as a result of which the following important conclusions were unanimously agreed upon:

1. That teachers' marks are ordinarily very apt to be inaccurate.
2. That, due to the personal standards of the teacher and to individual marking systems, the work of the same child may be graded very high by one teacher and very low by another.
3. That such irresponsible rating may work serious injustice to the children.
4. That the work of different schools cannot be accurately compared under such conditions.

And best of all, these conclusions were mostly arrived at by the teachers themselves with the help of a few leading questions on my part. Was it worth while? Would those teachers ever again correct test papers with the same self-satisfied assurance of the infallible justice of their marks? Would they ever again feel fully justified in retarding a pupil because he is given a mark of 68 when the passing mark is 70?

It was now time to bring the meeting to a close. Several wanted to know what better method of marking could be found. I advised all who were interested in solving the problem to order at once Monroe's book, "Measuring the Results of Teaching,"[2] and to think the matter over until our next meeting.

[2] Monroe, Walter S., *Measuring the Results of Teaching*. Boston: Houghton Mifflin Company, 1918.

The Second Meeting

In preparation for the next meeting I made a list of twelve history questions on the events leading up to the American Revolution. The questions were as follows:

1. How were laws made for the colonies?
2. What sort of governments had the colonies?
3. How was the commerce of the colonies regulated?
4. What kind of a king was George III?
5. What was the Stamp Act?
6. What were the objections to it?
7. What was the Stamp Act Congress and what did it do?
8. How did the British try to keep the colonists in order?
9. What friends had the colonists in Great Britain?
10. What was the Act of Association?
11. How was it enforced?
12. Distinguish between the real and the apparent reasons for the Revolution.

When we were again assembled I passed two copies of this list to each teacher with the request that they rank the questions one, two, three, etc., in the order of their difficulty beginning with the easiest, and that they write these ranks opposite the questions on one of the papers which they had received. When they had finished, I asked each teacher to read the ranks she had assigned to the questions beginning with question one, while I tabulated them on the board in such a way that vertical columns would show the various ranks assigned to the same question and horizontal rows the ranks assigned to all the questions by each teacher. I regret that I have lost my copy of the original table, but the general arrangement (for four teachers only) is shown below without the correct figures.

Teachers	Rank Assigned to Each Indicated Question											
	1	2	3	4	5	6	7	8	9	10	11	12
A	6	4	2	7	9	12	3	8	5	1	10	11
B	2	7	11	3	9	10	4	6	12	5	8	1
C	5	12	8	3	10	4	7	1	6	11	9	2
D	1	5	12	4	6	9	11	7	8	2	12	3

This tabulation showed such wide variations in the ranks assigned to the same question by different teachers, that a lively

interest was aroused. In the ensuing discussion further conclusions were reached which may be summed up as follows:

1. That if a test is to measure accurately a pupil's knowledge of the ground covered, the questions must be so selected as to bring out the most important ideas in the subject matter.
2. That there should be a large proportion of thought questions in order to call for something more than a parrot-like repetition of memorized facts, so that the better students may be enabled to demonstrate their superior reasoning abilities.
3. That some accurate method of weighting the questions must be used if accurate measurement of pupils' abilities and knowledge are to be expected.
4. That the weighting of questions by individual teachers on the basis of their personal judgments is mere guesswork since, as was shown by our experiment, they cannot even judge accurately the relative difficulty of questions.
5. That the average of the judgments of a number of teachers is apt to be more accurate than the judgment of a single one.

In line with this last conclusion we then proceeded to find the the average rank assigned each question and, with these averages as a basis, to rank the questions according to their relative difficulty as determined by the combined judgments of the teachers. Next I asked them to use these ranks as a guide and to assign values to each question so that the sum of the values would equal one hundred. It was observed that for each question the variations in values were not as wide as the variations in ranks had been. This was attributed to the fact that the order of difficulty used as a guide and the limit of one hundred placed on the sum of the values would necessarily permit much less variation. The results are shown in Table III.

The lesson of the day had been so well assimilated that someone suggested that since the ranking of the questions by each individual was all guesswork as proved by our experiment, then our averages were only averages of guesswork and hence likely to be far from reliable. For her part she did not see that we had any real proof that question twelve for instance was any harder than question three, and more than that she didn't see how we were going to prove that it was. Then came what I was hoping for but hardly expected.

"The eighth grades have recently completed their study of the period of history which the questions cover. Why not give them

TABLE III. WEIGHTS ASSIGNED TO TWELVE HISTORY QUESTIONS ACCORDING TO TEACHERS' JUDGMENTS

Question No.	Average Estimated Rank	Average Estimated Value
1	8	10
2	4	4
3	9	12
4	1	1
5	7	9
6	5	7
7	2	2
8	6	8
9	3	3
10	11	14
11	10	13
12	12	17 Total . . 100

as a test to these grades and so find out which questions are hardest?" This from a bright girl who had attended one summer session at Plymouth Normal.

"How could you prove which questions are hardest by doing that?" I asked.

"Why, if the questions were given to a large number of pupils, would it not be safe to conclude that the one which is missed the most times is hardest and that the one missed the next largest number of times is the next hardest, and so on?"

The practical side of this suggestion appealed to the majority of the teachers at once and they were eager to try it. Since this was the very thing I had intended to propose in order to get material for discussion at our next meeting, you may be sure that I was not slow in taking advantage of their enthusiasm. So it was decided to give this list of questions as a history test to all the eighth-grade pupils in the district. The teachers were to bring the corrected papers with them to the next meeting.

The Third Meeting

I watched the faces of the teachers as they gathered for our third meeting. The dazed look had passed from most of them by this time and a look of intelligent comprehension was taking its

place. The books which they had ordered at my request had arrived during the week and they had been reading them. They were apparently beginning to get their bearings. There was a business-like air in the way they took their seats and prepared for the afternoon's work that augured well for the future.

The meeting having been called to order we began to investigate the results of the history test. It had been given to 106 pupils. As I called a question by number, each teacher told me how many times her pupils had missed it. I wrote the figures on the board, a column for each question. When the results had all been read they were summed and tabulated as in Table IV.

TABLE IV. RANK OF HISTORY QUESTIONS IN DIFFICULTY, BASED ON NUMBER OF TIMES MISSED

Question No.	1	2	3	4	5	6	7	8	9	10	11	12
Times missed	28	44	40	8	23	20	35	29	16	60	56	51
Real rank	5	9	8	1	4	3	7	6	2	12	11	10

This table shows that question four is easiest since it was missed the least number of times and that question ten is hardest since it was missed the greatest number of times. Also, that question ten is seven and a half times as hard as question four because it was missed seven and a half times as often.[3] This would indicate that, in weighting the questions, number ten should be valued seven and a half times as much as number four. When these points had been discussed, even the teacher who was dissatisfied at our last meeting was convinced that we had found a fairly accurate method of evaluating questions.

The next step was to find values for the questions according to their relative difficulties. We first assigned to the easiest question (number four), which was missed eight times, a value of one. Since the next question in order of difficulty (number nine) was missed sixteen times, or just twice as often as number four, we gave it a value of two. Question six was missed twenty times. As twenty is two and a half times eight, this question was valued two and a half. The same method was followed throughout.

[3] Statistical difficulties involving the location of the zero point were ignored.

The sum of these values was found to be 51.5. Since the sum of the teachers' estimated values (Table III) was one hundred, it was necessary for comparative purposes that the sum of these values should also be one hundred. Evidently if each of them were multiplied by two, the sum of the resulting proportionate values would be 103. This was near enough to one hundred for our purposes and it had the advantage of yielding whole numbers. These may be seen in the last column of Table V, which summarizes the results of our study of the history questions.

TABLE V. SUMMARY OF RANKS AND VALUES FOR EACH OF TWELVE HISTORY QUESTIONS

Question No.	Estimated Rank	Real Rank	Times Missed	Estimated Value	Derived Value	The Same Multiplied by 2
1	8	5	28	10	3.5	7
2	4	9	44	4	5.5	11
3	9	8	40	12	5.0	10
4	1	1	8	1	1.0	2
5	7	4	23	9	3.0	6
6	5	3	20	7	2.5	5
7	2	7	35	2	4.5	9
8	6	6	29	8	3.5	7
9	3	2	16	3	2.0	4
10	11	12	60	14	7.5	15
11	10	11	56	13	7.0	14
12	12	10	51	17	6.5	13

After the teachers had rated a few history papers using the two sets of values, we concluded our experiments. Crude and inaccurate they were, of course; but they were not intended to contribute to the statistics of educational measurements. They were rather designed to exemplify to the teachers that spirit of scientific investigation which is so rapidly making over our school system, and in particular to convince them that teachers' ratings as ordinarily made are unreliable. And they had served well. All but two or three of the teachers had by this time expressed their firm conviction that tests devised by teachers and rated according to individual standards are of little use in finding out a pupil's real knowledge or ability, or in determining his standing with regard

to other pupils of his age and grade. These two or three teachers no amount of accumulated evidence could convince, because everything was unalterably settled to their way of thinking before the experiments were tried. The old methods had been in use for hundreds of years and must, therefore, be better than anything new. It is useless to waste time with such people. The only thing to do is to get rid of them at the first opportunity.

A few more points were cleared up by general discussion. Someone objected that, with a passing mark of seventy, it was evident from the scores that two-thirds of the pupils would fail to pass the test. This brought up an important point. Is it fair to mark pupils on a percentage basis with 100 standing for perfection? Some thought it was. Some thought it wasn't. Most hadn't thought about it, but since the matter was called to their attention they were inclined to think that it was hardly fair. What fairer way could I suggest?

I then explained what is meant by a median score, demonstrated with figures on the board the difference between median and average, pointed out the advantages of the median, and then proposed that we find the median score of the 106 history papers using the derived values for marking. This was found to be 65.1, to which I added 10 percent of itself and suggested that we use the result, 71.6, as a standard score. Any child who got this score would be given a mark of 100. That is, a pupil's mark would be the percent that his score was of the standard score. For instance if pupil A had a score of 26, his mark would be 26 divided by 71.6 or 36. If pupil B scored 63, his mark would be 63 divided by 71.6 or 90. A pupil who scored 85 would receive a mark of 118. In using letters for marks this pupil would be marked A+. It was pretty well agreed that this method of marking would be eminently fair, provided standard scores were available. The method may not be scientifically accurate; but it is certainly fairer than the ordinary method of marking on a percentage basis, and it has the advantage of simplicity.

Then someone wanted to know if we should do away with teachers' tests entirely and depend on the standardized tests altogether. We finally decided that teachers' tests should be used often by way of written reviews for the benefit of the pupils, but that the results should not have too much weight in determin-

ing the pupil's final standing. The pupils themselves, however, need not know how much or how little weight these tests might have on their grading for promotion.

Most of the standardized tests which I had ordered were at hand for this meeting. Samples of these were given to the teachers to study in connection with their new books on educational measurements. After explaining briefly the painstaking methods employed in deriving these tests and scales, I dismissed the meeting.

The Fourth Meeting

Our fourth meeting was devoted to the actual work of giving and scoring the tests. A few obvious facts were first emphasized, such as the need of accurate timing in cases where time was a factor, the fact that no help should be given the pupils other than clear and complete directions, and the fact that when directions accompanied the tests they should be followed explicitly. I also warned the teachers to beware of copying on the part of pupils; and I then proceeded to administer the tests to the teachers just as they should be administered in the classroom. As each test was finished we corrected and scored it, each teacher correcting and scoring her own paper for practice. In this way many obscure points were cleared up. Each teacher had seen each test given and had herself corrected and scored a sample of each test properly. They kept these corrected samples to use as models in correcting the tests which they were to give in their schools. Unless some such precaution is taken beforehand, the superintendent is sure to be surprised and dismayed when he gets the first packages of tests from his teachers. He will find that he must either return them to the teachers with time-consuming directions and explanations or go over most of them himself in order to get accurate results. I had been through such experiences before on a small scale and intended to avoid them this time as far as possible.

Since the tests were to be given to pupils the following week, each teacher then received a sufficient supply for her school. The tests were to be corrected and returned to me by mail within ten days. Unless a definite time limit is set and strictly adhered to, batches of tests from the various schools will come straggling in for two months after they are given out. Thus a few procrastinating teachers can delay the superintendent's part of the work to a

very annoying extent. In order to tabulate and study results for the district as a whole he must have all his data in hand at one time and as soon as possible after the tests are given out. In this case I was particularly anxious to have all the tests in on time as the grading of the schools was being delayed pending the results. So I laid special emphasis on the ten-day limit.

The teachers were further warned that, although I had no reason to distrust anybody, the matter was too important to permit taking any chances. Accordingly, I proposed to check the work of each teacher by giving one or two of the tests in her school *after* she had given all of them. By comparing the results of my tests with theirs of the same kind I could readily detect any gross carelessness or intentional dishonesty on the part of the teachers. There is considerable temptation for some shortsighted teachers who know that their own efficiency is being measured by these tests, to stretch the time limit or to give illegitimate aid to the pupils, or even to drill on the test itself, in the effort to make their classes show up well. Of course any intelligent teacher, knowing how the tests are used, would see the shortsightedness of such a policy since it is evident that what might be gained on one test would be lost on the next. However, two or three such dishonest teachers may, at first, work sad havoc with the accuracy of a superintendent's final figures regarding the efficiency of his schools and teachers unless he takes some precautions to discover the culprits at the start. Hence my method of checking their work.

Above all, tests should not be permitted to fall into the hands of pupils. Since they are to be used again and again to measure progress, teachers should be impressed with the importance of safeguarding them. As fast as the papers are finished they should be taken up to prevent possible copying of the questions or problems by the pupils in preparation for future tests. Of course, this does not apply in the case of the writing, spelling, and composition scales because of the different manner in which they are used. In fact it is well to display these scales on the walls of the school room so that the pupils may try to measure their own work.

This ends the account of how the cooperation of these teachers was secured in the use of standardized tests. It may not be intensely interesting in the reading, but it was assuredly interesting in the doing. Results since have proved its value. I can recommend the general procedure to anyone with a similar problem to solve.

FORMULATION OF METHOD IN READING[1]

H. A. Brown
State Normal School, Oshkosh, Wisconsin

The rapid accumulation of scientific data in education is making possible the formulation of a far more effective educational practice than has hitherto existed. In reading this is particularly true. The physiology and the psychology of reading have been investigated, so that one can now give a reasonably accurate scientific description of the act of reading. How the individual recognizes the printed symbols and how he assimilates the content underlying them can be stated with some degree of definiteness. The results of objective measurements of reading ability have thrown a great deal of light both on the reading act and the teaching process. In a word, a considerable body of facts has been accumulated in the last few years as the outcome of various types of investigations. The result is that we have a body of information sufficient to form the basis of a reasonably valid, scientific procedure in teaching.

Advance in teaching reading, however, has not kept pace with increased knowledge about reading. One sees even in the so-called best schools practices which can in no way be justified in the light of present knowledge. I am convinced that there is a very great need of a reformulation of practice on a somewhat higher plane. It will do little good to accumulate more data from measurements and investigations until better and wider use has been made of information now at hand. In making these statements, I am not unmindful of the great progress which has been made during the last ten years, and especially during the last five years, in the teaching of reading.

Of course, the limits of this address would preclude a complete discussion and reformulation of procedure in teaching reading. It is my purpose, however, to indicate some lines along which this reorganization should take place, and to suggest certain definite steps in the direction of improved methods of instruction which need to be taken immediately, in order to bring practice more

[1] An address delivered before the National Association of Directors of Educational Research, at Cleveland, Ohio, February 26, 1920.

nearly into harmony with the enlarged body of scientific information which has been brought forward by investigators in the last few years.

Method of Approach to the Problem

The best approach to this problem of the pedagogy of reading must involve three distinct lines of investigation.

1. The first important step is a study of those fundamental and significant facts of the physiology and psychology of reading, which throw light upon the processes involved in the act of reading.

2. From a study of all the known facts at hand in this field, certain hypotheses or tentative principles of teaching may be established, and we may then proceed to teach reading in accordance with the principles thus established. In a word, on the basis of the data from the physiology and psychology of reading, it is possible to develop an experimental technic of instruction for practical use.

3. Before we can be assured that our method is correct, it must be evaluated in terms of the ability of the children to read. Thus elements of efficiency and points of weakness will be discovered. A process by which the former are selected and retained and the latter eliminated will ensue, until a pedagogy of reading has been evolved, which will be in harmony with the laws of the most economical learning.

This, in brief, is the method which must be pursued not only in reading but in all school subjects in order to advance toward better methods of instruction. As the body of significant facts concerning the mental processes of children and the nature of the learning process accumulates, this gradual rebuilding of educational practice becomes one of the most important problems of the day.

Definition of Reading

Preliminary to a discussion of classroom technic in reading, I should like to state definitely what my conception of reading is. In a word, I desire to define reading in a way which will indicate what the essential factors in the process are, from my point of view.

To read, as the term is here used, is to interpret, organize, remember, and deliver the thought conveyed by the symbols

of the printed page. This statement indicates that there are three main factors involved in effective reading ability. In order to be able to read at all, the individual must interpret the content underlying the printed words on the page. But to understand only is not enough. It will be of little benefit to the pupil merely to comprehend his history lesson as he reads it over. If later, when the book is closed and he is in the class, he can remember little or nothing of what he read, he will not be able to participate effectively in either recitation or discussion. Therefore, his reading will not be effective unless, in addition to interpreting and assimilating ideas, he can also recall them. To carry the thought one step further, his reading cannot be said to be effective unless he can in some way return the ideas which he has assimilated. This may take the form of reproduction, or of using these ideas in some appropriate way. Therefore, effective reading involves carrying the thread of the story, argument, or discussion, grasping all of the essentials, sensing the significance of the whole, retaining it in the mind in well-organized form, with the ideas in rational sequence, so that they can be delivered in some form. According to our conception, this is real reading. Ability to read and ability to study are, then, in part synonymous terms, although study ability of course includes more than reading ability.

What I have defined above is the complete act of reading, and I am quite aware that it probably ought not to be called reading. I realize that in many cases interpretation of what is read is entirely sufficient. Today, for example, I desired to go from one part of the city to another. With printed directions in my hand, all I needed to be able to do was to understand the directions. However, there are many cases in school and in life, in which ability in what I have called the complete act of reading has so great value that it needs to be taught to all children. In fact it is an indispensable ability.

Factors in Reading Efficiency

In determining reading efficiency, two factors must be taken into consideration. In the first place, of two readers whose grasp of content is equally great, that one who can read the larger amount in a given time has the greater reading ability. Rate of reading, then, is one of the two factors. In the second place, of

two people who can read the same amount in a given time, that one whose grasp of content is greater is the better reader. In this case the individual who can reproduce the larger amount of what he has read, and who can do so the more correctly is the more efficient. The two things, then, which enter into a complete measure of reading ability are: (1) rate of reading; and (2) degree of effective comprehension (including ability to interpret, organize, remember, and deliver in some form).

Now, the main purpose in teaching children to read is the development of power to get thought from the printed or written page. The fact that a child or a class has the ability to pronounce words fluently cannot always be taken as an indication of good reading, for it may not be accompanied by adequate comprehension of the thought. For example, a child may be obliged to spend a half hour to read a page so that he comprehends the thought sufficiently to reproduce 80 percent of the content. Another child may be able to read the same page with equal understanding in a few minutes. The second pupil is, of course, a much more efficient reader. Real efficiency demands a satisfactorily rapid rate of reading combined with an adequate power of comprehension. In the measurement of reading ability it is desirable to combine these two factors into a final measure which shall represent the total reading ability of a child or a class. Perhaps it will prove to be impossible to do this. Reading ability may be two-dimensional.

Initial Steps in Teaching Reading

The reading lessons of the first three months of school are all-important. They may be so conducted as to facilitate or hinder economical learning during the following years. The classroom technic is important. We now know from investigations that reading is a complex, muscular, photographic, and assimilative process. There is every reason to believe that the act of reading is subject to the laws of habit formation. If the teacher's procedure inflicts upon the child during the first three months habits unfavorable to effective reading ability, they will be difficult to break in later years.

Not long ago I was familiar with a class which was taught reading during the first year by the alphabet method. The children

were taught to spell each unfamiliar or difficult word. This resulted in a pronounced habit of attention to textual details. At the beginning of the second year a complete change in method was adopted and every device known to modern educational practice was used to break the bad habits formed during the first year. Rapid silent reading from books, quick recognition devices for reading the meaning of words and phrases, and an abundance of silent reading both in school and out were persistently used. At the end of the fourth school year, however, not one of these children had ceased to pause and silently spell each difficult word. They seemed incapable of visualizing a new word quickly and accurately on account of the early habit which appeared to persist in spite of all efforts to break it. This is not an uncommon experience on the part of teachers. May it not be, then, that when an early habit of reading with attention to textual details has been established it is impossible ever to eradicate it completely? May it not be that during the first year either readers are made or bad habits are permanently developed? If this is true, it is highly important that a technic be established for the first steps in reading, whereby efficient habits of word recognition and of assimilation of content may be provided.

I believe that some of our practices should be radically changed in some important particulars, and that real reading must be secured from the outset. The good reader in perusing the printed page, while words are continually before his eyes, does not really see the words. He is engaged primarily in thinking meanings, in comprehending the thought. The early lessons in reading should produce this effect. It is my conviction that most of our methods are seriously at fault, because they produce a consciousness of word forms which is both unnecessary and highly unfavorable to the development of efficient reading habits.

How, then, may the first steps in reading be taught so as to reduce word consciousness to a minimum? In answering this question I shall describe the procedure which is used in the training department of the State Normal School at Oshkosh.[2]

[2] Material assistance has been rendered in preparing this section by Miss Ruberta N. Smith, head of the department for training primary teachers, at the State Normal School at Oshkosh. Several of the following paragraphs have been written by her.

The First Lessons

The earliest reading lessons involve only words which are already a part of the children's oral vocabularies and which are clearly associated with familiar objects, ideas, or experiences. The teacher's problem is to make associations between the known auditory forms and the unfamiliar visual forms.

Reading in the first grade does not begin for some time after school opens. Frequent periods, however, are devoted to story telling, simple dramatization, and conversation about interesting topics. On these topics the children are led to express themselves freely and spontaneously. A basis is thus laid for the first reading lessons. This goes on for days and is guided by the teacher so that the words, phrases, and sentences on which she plans to base the first reading lessons are used again and again by both the children and herself.

After a time short sentences, the auditory forms of which are familiar to the children, begin to be introduced. For example, perhaps some of the children have pet rabbits, and on a certain day these are brought to school and become the center of interest. The next morning a life-like picture of the mother rabbit and the young rabbit may appear on the blackboard with the sentences *Here is big mother rabbit* and *This is little baby rabbit.* On the previous day these two sentences were used purposely by teacher and children, and they are now used again many times. The visual forms of these sentences, seen in connection with the objects and in connection with the frequent oral repetition of the sentences, quickly become associated with the objects and the auditory forms. These sentences are to the children unanalyzed, visual wholes representing certain meanings, just as spoken sentences are to young children complexes of sounds conveying certain meanings which they grasp as unanalyzed wholes. Without carrying the illustration further, I may say that many short sentences are made meaningful to the children in this way as unanalyzed visual wholes, in a more or less incidental manner, before any reading is attempted.

In presenting the first reading lesson the teacher leads the children into conversation on some topic about which they have talked on previous days, at the outset arousing the interest and securing the attention of the class. When the latter is at its maximum, the

teacher writes a sentence of from two to four words on the blackboard, after she and the pupils have used it orally a number of times. This sentence is at the center of interest in the conversation. Both teacher and children will naturally have occasion to repeat it several times. During the conversation, and at a point where the sentence would naturally intervene, it is written by the teacher as she utters it, the visual form thus appearing while the auditory form and the thought are in the attention of the children. As the sentence is repeated, the teacher points to it so that the children realize that the particular visual complex on the board stands for the particular meaning in question.

The sentence is then read by various pupils until all the children in the class can react to it as a whole without hesitation. Two sentences closely connected in thought and requiring as few different words as possible are treated in this manner in the first lesson. Since the children are not focally and explicitly conscious of word forms, much less of letters or phonetic elements, a right start in reading has been made.

On the second day, the sentences of the first day are reread several times and one more sentence is developed, in the same manner as were those of the previous day. This continues for a number of days, not more than three new sentences being introduced in two days. Each day the sentences previously developed are reread several times, and since they form a connected story, this rereading is always interesting. It is held to be excellent practice for the children to react in this way again and again to an entire sentence as a whole. They soon catch the idea and when a sentence is pointed to, they respond with the whole thought as quickly as in their own animated conversation. To get established the habit of reading entire sentences as wholes, in the same connected manner which children use in conversation, with the thought focal in attention, is the fundamental object at the outset in reading. In order to accomplish this, one can afford to make haste slowly.

After a number of lessons the children have a sufficient vocabulary so that for two or three days new sentences may be constructed with the words already familiar. Whole sentences continue to be the basis of instruction. The object is to read as much as possible with the small vocabulary. Words must be

repeated again and again, but always in new combinations, and in connection with a content of vital interest to the children.

Blackboard work continues for six weeks or two months. It furnishes opportunity for visualizing sentences with a very brief exposure. A shade hung at the top of the board is a useful piece of apparatus. Sometimes sentences may be written on strips of cardboard and briefly exposed. In none of this work are the children ever allowed to read sentences word by word.

From the very first, these lessons consist of connected stories. Even a reading lesson of a half-dozen sentences is in the form of a story about something which will grip the interest of the children and hold their attention. It ends in a climax which leaves them with a desire for further reading. Especially does the teacher seek to avoid the kind of disconnected, disjoined sentences often found in primers.

Reading from Books

At the end of about two months of blackboard work, the class begins to read from an easy primer. The transition from blackboard script to the printed page is usually accomplished with very little difficulty.

The reading continues to follow quick perception lines. Slow rendering of the thought word by word is always avoided. For example, each child may be provided with a strip of cardboard which may be moved down to expose a single line or sentence and, then, at the word of the teacher, quickly moved back, thus permitting only a glance at the line. The children should at all times get the entire thought of a sentence before they are permitted to read it orally. It is not recommended, however, that children be required to look away from the book or at the teacher all the time they are giving the sentences.

It is not necessary or desirable to read one primer through before leaving it. The first ten or twelve pages of a half-dozen or more primers should be read first. Later in the year these books may be finished and several children's classics may be read. The main purpose is to read as extensively as possible from easy and interesting material.

Development of New Words

The matter of word development is of the greatest importance in the early stages of reading. If the association between the visual form and idea is faintly made, the visual form will recall the idea faintly or not at all.

During the earliest stages of reading each new word, as has been previously suggested, is carefully developed through pictures, objects, and conversation, and presented in sentences. By the middle of the year, however, when the children are reading rather extensively from books, it is not necessary to develop all new words previous to reading. When an unfamiliar word occurs the teacher pronounces it, if the child who is reading is unable to do so. He pronounces it after her and the reading continues without interruption. The teacher notes a few of these words. After class she writes them on the board and devotes two or three minutes to practice in pronouncing them as unanalyzed wholes. The teacher pronounces each word slowly, clearly, and distinctly as a whole and the children pronounce after her in the same way. The period ends with a little practice in pronouncing the various words quickly as the teacher rapidly points to them. With a little instruction by the teacher, children thus catch the knack of pronouncing words merely by practice in pronouncing them. Experience indicates that children unconsciously work out their phonetics in this manner without such a large amount of abstract drill as is sometimes given. A very large amount of time and energy is thus saved.

Visual Analysis of Words

The word-pronunciation exercises of the first year are continued in the second year, and during the latter part of the year they take the form of practice in the visual analysis of words. This is conducted in the same general manner as that suggested for the exercises in word pronunciation. At the close of each reading lesson the teacher places upon the board, as previously, a half-dozen or more of the difficult words of the lesson. She indicates the syllabic form of each, pronounces it distinctly, and has the pupils do so. After the words have thus been gone over several times, the teacher points rapidly to the different words and has them pronounced quickly as wholes. After about two

weeks of this kind of work, the pupils begin to have practice in dividing into syllables and pronouncing words taken from the reading lessons and written on the board. Practice is given in quickly glancing at words on the board followed by syllabication and pronunciation. Children are encouraged to pronounce unfamiliar words at sight as they occur in the reading. By the time they are well started in the third grade they are able to syllabicate and pronounce any ordinary word met in their reading.

Not over four or five minutes a day are devoted to the exercises in visual analysis, which serves as the rounding up of the practice in word pronunciation of the first year and a half.

Drill on Words and Word Groups

Children must be taught to grasp not only words but also word groups as wholes. The ability to read word groups by merely noting some familiar characteristic contributes greatly to rapid reading. To develop this ability drill is given with cards containing words and word groups of frequent occurrence. This drill begins after some facility in reading has been acquired, not earlier than near the end of the first year.

The drill is rapid and the periods brief. As nearly instantaneous perception as possible is required. The technic of the drill is such that the card is raised to the proper position for the recognition and then lowered in a way which will provide an exposure sufficient for one brief glance with the card at complete rest. During its movement into position and back the words must not be exposed.

Phonetic Drill

Certain considerations must be taken into account in connection with the problem of phonetic drill. What reading habits are most desirable? What procedure in teaching is conducive to their development? These are questions which must receive serious consideration. Speed is a highly desirable factor in reading ability, merely as a timesaver if for no other reason. All habits which lead to slow and ponderous reading must be avoided. Speed in reading depends in part upon the size of the units of recognition, i.e., the amount which can be seen by the reader during eye pauses, and in part upon the duration of eye pauses.

The individual who sees only a small portion of the line during an eye pause must make many fixations. The longer the eye pauses the slower will be the reading, other things being equal. Will not a habit of attention to textual details inevitably result in slowness in reading? Will not a habit of recognition through dominant characteristics of words and phrases, with their minute particulars marginal in attention, become a distinct aid in interpreting thought? Is not the inevitable result of two or three years of persistent drill on the isolated elements of words bound to be a predominant habit of attention to the minute particulars of words? Can the eye be trained to take in large units by practice on small ones? Will not prolonged attention to word forms effectively prevent the development of a habit of thinking meanings as one reads?

In view of the facts at hand, it seems to me to be a radically wrong procedure to introduce such a large amount of abstract phonetic drill in the earliest stages of reading as is often found. It seems to me to be a very obvious inference that the predominant emphasis in method at the beginning must be upon the instantaneous recognition of words and word groups, the mental process being concerned with meanings rather than with word forms.

Is not the child afforded the best kind of drill on phonetic elements when he is made to react to them constantly in reading sentences containing words in which they occur? May not the additional time usually devoted to abstract phonetic drill profitably be given to more reading? Will not children thus acquire a feeling for sound and an unconscious mastery of phonetics which, in the end, will give them sufficient power over new words and avoid developing the consciousness of words which prolonged attention to their minute particulars is likely to produce?

It seems to me that here is a definite problem which should be subjected immediately to scientific investigation. It will be futile, however, in the comparison of reading methods to test groups of children taught by different methods and draw conclusions from the results without taking into consideration a variety of other factors. For example, the time elements must be the same in both groups and the amount of supplementary outside reading must be the same. At least one investigation of this problem has fallen far short of securing a valid result on account of plain

disregard of important factors of this kind. A carefully controlled experiment with every other factor identical in two groups of children with the exception of the presence of phonetic drill in one and its absence in the other would yield a most interesting result. At present, I believe the data at hand raise serious doubt as to the value of a large amount of abstract, phonetic drill during the earliest stages of reading instruction.

Experience with One Class

The children of the third grade in the training department of the State Normal School at Oshkosh spent the first two years in extensive reading of simple stories for pure enjoyment, with their attention entirely on the content. During the reading period all words not instantly recognized were told immediately. Word development was taken up in separate periods, the words being incorporated in phrases for drill, and no attention being called to the elements of words. Approximately fifty minutes per day were spent on all phases of reading, including silent reading, word development and study, and oral reading. Silent reading was begun early in the first year. Before the end of the second year the children began to pronounce new words without help, applying independently the general knowledge of phonic laws acquired unconsciously through much reading.

Beginning with the third year, the children were encouraged to organize the phonic facts so gained and the four or five most common principles of phonics were taught. A short period at the end of the day was devoted to *pronouncing long words*, as the children called it. This period became so popular that only with difficulty were the children persuaded to omit it on special occasions. Recently they were pronouncing with ease and enjoyment such words as hydrocarbon, hydrometer, tuberculosis, capitulate, centigrade, epidote, actuate, and many others of equal difficulty, without the aid of diacritical marks or context.

These children have developed insatiable appetites for reading, and they read intelligently and accurately with good expression and excellent speed. Of the two daily reading lessons in the third grade, one is devoted to oral and the other to silent reading. The children are capable of reading any material suited to their age in content, and they read great numbers of books during spare periods and out of school hours.

Reading Ability as a Phase of Study Ability

Definite training in silent reading is needed for the purpose of developing in children the ability to study efficiently. This training should begin early in the first year and continue until a mastery of the printed page has been developed. Schools need to take very pronounced steps in the direction of developing a special technic for teaching silent reading.

Not long ago I observed a lesson which seemed to me particularly worth while for intermediate grades. While visiting schools in a certain community, I stepped into a room in which were third and fourth grades. I observed that the teacher was asking questions from a book. The children appeared to be reading silently and rather rapidly. Suddenly the teacher clapped her hands. The children ceased reading immediately and closed their books. Various children were then called upon to tell the substance of what they had read. They appeared to me to have a remarkable ability to reproduce in great detail. I remained in the room, and when the reading began again it was conducted in the following manner. A question was asked by the teacher to which the first two or three sentences of the paragraph furnished the answer. The children read rapidly to get the facts. By the time they had about grasped the meaning of these sentences another question came from the teacher, and the process was repeated. In this manner there was a series of rapid-fire questions from the teacher and very rapid silent reading on the part of the class. In this way the class covered in a few minutes about two pages before the teacher again clapped her hands as a signal for the closing of the books. Again there was very efficient oral reproduction of what had been read. I learned that the reading was entirely at sight.

I remained throughout the period, which lasted about fifteen minutes. It was the best example of competent teaching of silent reading which I have ever seen. By her manner of questioning the teacher forced the class to read at their maximum ability for a short period. Five times as much was read as is ordinarily read in an oral reading lesson. It should be pointed out that it is utterly futile to attempt to teach silent reading, from the point of view of developing study ability, by letting the pupils dawdle along at their own pace.

Silent reading should begin early in the first year. In its simplest form, it consists in following directions such as *Bounce the ball*, or *Run to the door*—these directions being developed by the class as a game. A little later a thoroughly familiar reading vocabulary may be used and a group of connected sentences presented one after another on the blackboard. The response to this type of lesson may be discussion, dramatization, or illustration, with the use of crayons, scissors, paper, clay, sand table, or blackboard. For example, a group of half a dozen sentences may describe some adventure of a kitten. After the reading has taken place silently sentence by sentence, the children may play the story or discuss the result of the adventure or its effect on the kitten, or spend a seat period illustrating the points made in the story.

More formal silent reading may begin in the second half of the first year. Children should now be able to grasp short sentences by quickly glancing through them, sensing the meaning without being explicitly conscious of the minute particulars of word forms. Good teachers will adopt devices that give the class but a moment for glancing through a sentence. The class may read silently a group of a dozen or more sentences, after which different pupils may reproduce orally the meaning of what was read. This is the beginning of teaching to read silently and of teaching to study.

It is held that the ordinary oral reading lesson is unproductive in grades beyond the third. Instead of this the reading may well consist of a relatively small amount of motivated oral reading and a large amount of silent reading for the purpose of teaching pupils how to study efficiently. The regular reading period as a separate part of the program may be dispensed with and silent reading may be taught in connection with other subjects. This will result in a considerable saving of time. For example, on one day the reading lesson may consist of silent study of a history assignment and may be conducted on history time. On another day geography may furnish the basis for a silent reading lesson. Again, on a third day civics may serve as material for reading. Each of these lessons will afford definite training in silent reading from the point of view of teaching how to study. Papers and special reports may be prepared from time to time. The reading of a report to the class is excellent motivated oral reading. Per-

haps a pupil may find at home a book which contains material not found in any of the books at school. Such a pupil will delight in bringing this book to school and reading from it to the rest of the class. The situation is thoroughly motivated, because he knows something which the rest do not know and therefore has a definite message to convey to them. This provides an incentive for him to read as well as he can. The audience desires to get the information and has an incentive to listen.

Without going into further detail it is sufficient to say that the resourceful teacher will have no difficulty in utilizing all of the different subjects of the curriculum as valuable material for training in silent reading. There will be no difficulty in finding plenty of situations for motivated oral reading. Whenever a pupil reads silently with a definite purpose to get information, or when he reads orally with a definite aim to convey information to others, either type of reading is properly motivated. It is doubtful if any other kind of reading is justifiable. It is also to be seriously questioned whether a pupil profits greatly from any other kind of reading.

Testing and Diagnosing Reading Ability

Although much progress has lately been made in measuring ability to read, all that has been done constitutes but a good beginning. One of the great needs in the measurement of reading ability is that tests be devised for determining levels of study ability. I have contended that the best kind of material for a reading test is a connected passage in the form of a short story, a description, a piece of exposition, or some similar kind of writing in which the child's ability to react to the whole selection or to as much of it as he can read in the time given is tested. We greatly need as test material connected passages of various known degrees of difficulty. Such material would not be especially difficult to prepare, and with it we could apply to a given child or class several tests, each more difficult than the previous one, until the exact level of study ability of the child or the class had been determined. It would be necessary to have several alternative tests for each degree of difficulty, so that a class or a pupil could be tested at frequent intervals.

Such a series of tests as this would prove of inestimable value in the upper grades and high school for the purpose of placing

pupils of approximately equal study ability in the same groups. One of the greatest sources of waste in teaching reading has been in keeping children under instruction in a given class after they have advanced to a level of reading ability far beyond that contemplated by the instruction. Probably great harm has been done such children by keeping them dawdling along at a pace far below their ability. On the other hand, there are pupils in every class in the upper grades and high school, whose level of study ability is so much below that contemplated by the books in use and the instruction given, that it is impossible for such pupils to succeed. A reading test designed for the purpose of diagnosing levels of study ability, together with instruction in silent reading, would prevent many of these pupils from becoming failures and would remove a source of waste and inefficiency from the schools.

THE CORRELATION RATIO

LEONARD P. AYRES
Russell Sage Foundation

In the March, April, and May numbers of the JOURNAL OF EDUCATIONAL RESEARCH there have appeared articles by the writer describing methods for computing the coefficients of correlation and regression that are shorter and simpler than those that have heretofore been in use.

The purpose of this article is to describe a short method for computing the correlation ratio which is a measure proposed by Pearson for computing correlation in distributions in which the relationship between the two series is a curve relation rather than rectilinear. Such relationships are frequent.

For example there may be a real relationship between the amount of money spent on municipal sanitation and the decline in the death rate. In extreme cases, such as have recently existed in tropical America, conditions have been found such that slight expenditures for sanitation resulted in marked reductions in sickness. As improvements go on, however, it is found that the expenditures must be greatly increased to effect even small decreases in the death rate. Similar conditions are common in the field of education as, for example, in the relation between the amount of time devoted to the practice of penmanship and the quality of the results obtained.

The coefficient of correlation will not accurately measure these curvilinear relations and the correlation ratio is the device that is proposed to obtain a trustworthy measurement of them.

The correlation ratio can only be computed for tables of distribution. It is the ratio between two standard deviations. One of these is the standard deviation of all the values of one series—say the subject series. The other standard deviation is that of the averages of the columns of the subject values from the average of all the subject values in the table. If the relationship of the two series is rectilinear the correlation ratio is equal to the coefficient of correlation. If it is curvilinear the ratio will be greater than the coefficient of correlation. It is always positive and never exceeds 1. Its symbol is η (eta). The best description of

its derivation and use is found in Rugg's *Statistical Methods Applied to Education*.

The computation of this measure may be illustrated by finding both the coefficient of correlation and the correlation ratio for a simple distribution table in which the relationship is roughly curvilinear. The coefficient of correlation is found by the formula:

$$r = \frac{\text{Sum of products of subject and relative items minus av. of subject items} \times \text{tot. of relative items}}{\sqrt{\text{Sum of squares of sub. items minus av.} \times \text{total of these items} \times \text{Sum of squares of rel. items minus av.} \times \text{total of these items}}}$$

The computations are shown in the accompanying tabulation in which "S," "R," and "T" stand for "subject," "relative," and "total" respectively. The letter "f" means "frequencies" and refers to the entries in the body of the table. The figures in the columns and rows at the right and bottom of the table give the operations necessary for finding the coefficient of correlation, with the exception of the two lowest rows at the bottom which are separated from the others and which represent the operations necessary for finding the correlation ratio.

As the tabulation shows, the coefficient of correlation is 0.73. The next step is to find the correlation ratio. This is the ratio between two standard deviations of which the first is that of all the items of the subject series. Since the standard deviation is the square root of the average of the squares of the deviations of the numbers of a series away from their own average, the first step is to find the sum of the squares of the deviations in the subject series.

This is 62 and it has already been found in computing the coefficient of correlation. It is equal to the sum of the squares of the subject series (678) minus the total times the average of that series [$678 - (136 \times 4.53) = 62$].

The next step is to find the sum of the squares of the deviations of the averages of each column from the average of the entire subject series. This work can be greatly shortened by following the method utilized in finding the coefficient of correlation. That

Subject Values	Relative Values 1	2	3	4	5	T	ST	SST
2	3	1	.....	.....	.....	4	8	16
3	3	1	.....	.....	.....	4	12	36
4	3	2	.....	.....	.....	5	20	80
5	1	3	2	.....	.....	6	30	150
6	.....	2	3	3	3	11	66	396
T	10	9	5	3	3	30	136	678
RT	10	18	15	12	15	70	Av. of sub. items, $\frac{136}{30}=4.53$	
RRT	10	36	45	48	75	214		
ΣSf	32	40	28	18	18		Av. of rel. items,	
R(ΣSf)	32	80	84	72	90	358	$\frac{70}{30}=2.33$	
$(\Sigma Sf)^2$	1,024	1,600	784	324	324			
$\frac{(\Sigma Sf)^2}{T}$	102.4	177.8	156.8	108	108	653		

$$70\times4.53=317$$
$$136\times4.53=616$$
$$70\times2.33=163$$

$$358-317=41$$
$$678-616=62$$
$$214-163=51$$

$$r=\frac{41}{\sqrt{62\times51}}=\frac{41}{\sqrt{3,162}}=\frac{41}{56.23}=0.73$$

method is based on the principle that the sum of the squares of the deviations of a series away from its average is equal to the sum of the squares of the numbers in the series, minus the product of the total of the series and its average.

Following the same principle we find the sum of the squares of the deviations of the averages of the columns away from the average of the whole subject series. The first step is to find for each column the square of its average multiplied by, or weighted

by, the number of cases in the column. In symbols the desired figure for each column is equal to $\frac{\Sigma Sf}{n} \times \frac{\Sigma Sf}{n} \times \frac{n}{1}$, which is equal to $\frac{(\Sigma Sf)^2}{n}$.

In finding the coefficient of correlation we have already found ΣSf for each column, so the new operations may be performed by squaring these numbers and dividing the result by the number of cases. This is done in the two added rows below the table. Thus, in the third column of the table the entries are 2 and 3 and they correspond to the values 5 and 6. The sum of the subject values times their frequencies (ΣSf) is then (2×5) plus (3×6) or 28. In the next to the last row this number is squared and in the last it is divided by the number of cases (5) giving a result of 156.8, which is the square of the average of the values of the column weighted by the number of cases. This may be verified by noting that the sum of the values of the column is (2×5) plus (3×6) = 28 and their average 28/5. This squared is 31.36, which weighted by the number of cases (5) equals 156.8.

The total of this last row of weighted squares of averages is 653. From this we get the sum of the squares of the deviations from the averages of all the subject values in the table by subtracting the product of the sum of the subject series and its average or (136×4.53 = 616) which gives 37.

Probably the best way for the reader to convince himself that this method is valid and to find out why it is valid is to experiment with simple illustrative cases, carrying the work through by the standard method as described by Rugg and by this method.

We now have two numbers, 62 and 37, which are the sums of squares of deviations and from them we must find the two standard deviations and the ratio between them which will be the correlation ratio. The work is as follows:

$$\eta = \frac{\sqrt{\frac{37}{30}}}{\sqrt{\frac{62}{30}}} = \frac{\sqrt{37}}{\sqrt{62}} = \frac{6.083}{7.874} = 0.77$$

Instead of dividing 37 and 62 by 30 before finding the square roots it is better to eliminate the two 30's. This shortens the work and does not change the relationship between the numerator and denominator.

The correlation ratio is 0.77 whereas the coefficient of correlation is 0.73. The difference is frequently much greater than it is here. The added work necessary to find the correlation ratio is slight. It consists for the most part in the two added lines of figures at the bottom of the table which are easy to get because they are derived from the other computations of the rows above them which have already been used in computing the coefficient of correlation.

There are two correlation ratios for each table of distribution, just as there are two coefficients of regression for each coefficient of correlation. In the case here given the other correlation ratio would be found by working with the relative series and the rows of numbers in the original table instead of with the subject numbers and the columns. The added columns of figures would be at the right of the main table instead of at the bottom. Generally the two correlation ratios are nearly identical so that it is not necessary to find both of them.

Following the method described, the formula for the correlation ratio becomes

$$\eta = \frac{\sigma_{AS}}{\sigma_S}$$

in which the subscript AS indicates that the numerator of the fraction is the standard deviation of the averages of the subject values in each column from the average of all the subject values, and the subscript S indicates that the denominator is the standard deviation of all the subject values in the table.

PUBLICITY CAMPAIGNS FOR BETTER SCHOOL SUPPORT

CARTER ALEXANDER
AND
W. W. THEISEN
State Department of Public Instruction, Madison, Wisconsin

I. THE PROBLEM

The most casual observer of the efforts to secure better school support in this critical period cannot but note the emphasis placed upon publicity campaigns or "drives." This emphasis is no accidental or temporary affair. In the first place, the public schools have been forced to compete with ever increasing taxes for other purposes. Demands for better highways, improved sanitation, increased protection to life and property, and increased state and governmental service in various directions have all tended to raise taxes and to make it relatively more difficult to secure the necessary school expenditures. In the second place, the standards of school service of today are much higher than formerly. Improved school buildings, more highly trained teachers, enriched curricula and a longer school term require more money than did the school programs of a few decades ago. Again, during the war period citizens became so accustomed to drives for large sums of money that any cause which does not now conduct a drive is almost certain to seem trivial or unimportant to them. Increased taxes and rising costs of living have combined to make taxpayers even more reluctant to increase school expenditures. In many places schools are now running up to the full tax levies which are legally possible without a direct vote of the people for an increase. There is probably not one school system in a hundred where the people can be induced to vote increased school taxes for the amount now needed, without a well-directed and vigorous campaign of publicity.

Such being the case, our problem is to discover the principles and procedures underlying the organization and operation of successful school publicity campaigns. How can their success be insured? More particularly, what should be the form of organization? How ought the constituency to be analyzed for points

of possible contact? What avenues and instruments of approach to the public can be utilized? How should arguments and illustrations that will really convince the voters be selected? How should publicity material be prepared and circulated to be most effective? These are questions which this investigation seeks to answer.

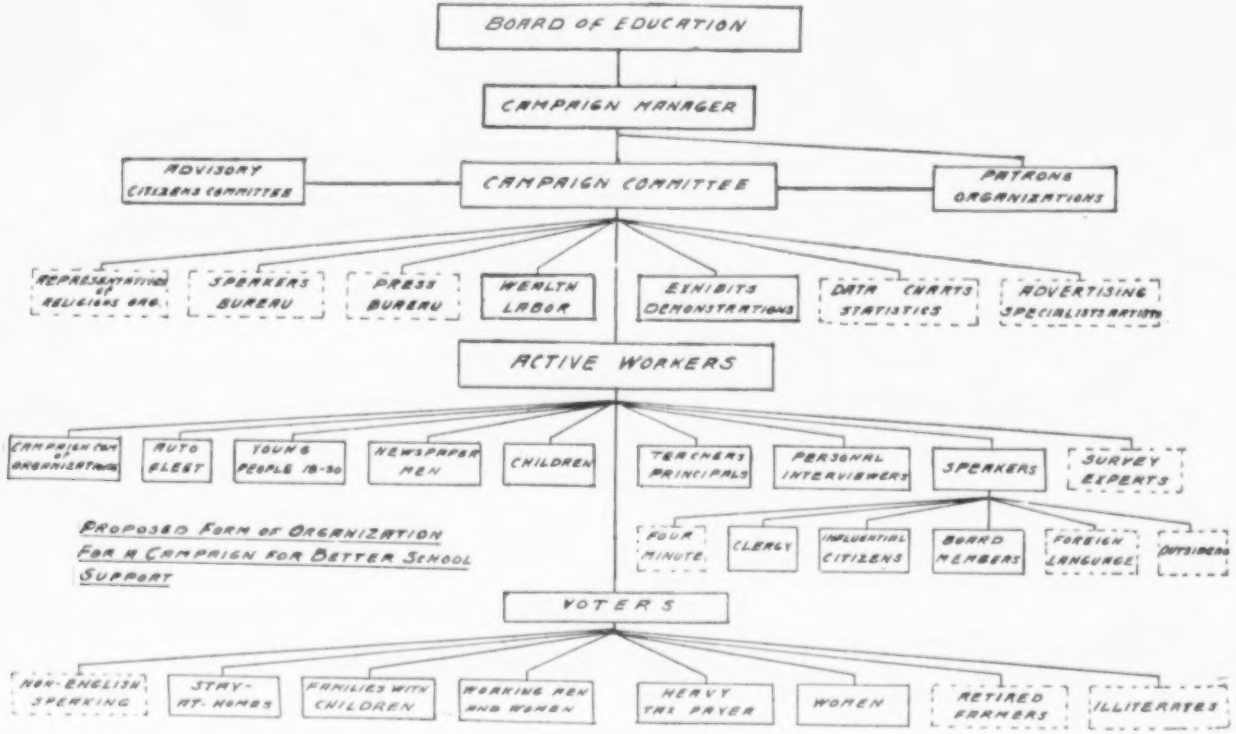
II. The Method

In this study we have attempted to follow the accepted procedure for matters involving as much social psychology as do publicity campaigns. This method is to analyze the work of successful men with a view to discovering the secrets of their successes and to formulating rules for achieving similar ones by other men. Accordingly, we have sought the solution of our problem in a careful study of school campaigns that have actually succeeded. There is an art, a skill in generalship, an indefinable something in leaders in school campaigns, which needs to be carefully analyzed and made available for other school men.

To secure accounts of publicity campaigns, a letter was addressed to superintendents in cities in various parts of the United States wherever it could be learned that such methods had been successfully employed.[1] The letter, which briefly explained the purpose of the study, contained a request for a careful description of the methods employed and copies of material used. A suggestive list of typical situations on which information regarding publicity methods was being sought was inclosed with the letter.

More than seventy superintendents representing various parts of the country and cities of all sizes responded generously. Most of them described the significant phases of the campaign in their own community and many sent printed or mimeographed materials that had been used. In several instances follow-up letters were written to secure specific details. Wherever opportunity was afforded through educational gatherings of various kinds personal interviews were held. The data thus gathered represented every general condition. The writers have studied the

[1] In this connection the writers are indebted to several members of the National Association of Directors of Educational Research for supplying information as to cities in their own states that had recently passed through a campaign. Dr. B. R. Buckingham furnished data for Illinois, Dr. E. J. Ashbaugh for Iowa, Dean F. J. Kelly for Kansas, and Dr. W. W. Phelan for Oklahoma.



PROPOSED FORM OF ORGANIZATION FOR A CAMPAIGN FOR BETTER SCHOOL SUPPORT

accumulated evidence with considerable care and from it have reached the conclusions set forth in this article.

III. The Results

In a study of this nature it is impossible to present the results in the ways that would be possible in a purely statistical investigation. Our presentation will therefore be of a descriptive nature. Again the bare results without the illustrations and data from which they are derived cannot carry the whole story. For that reason it has seemed advisable in this article to present only a brief sketch of the findings and conclusions. We have included a diagrammatic representation of the form of organization desirable for a successful publicity campaign, a composite of the best features of all the campaigns studied, and brief notes on some additional aspects of the problem of securing better school support. The complete results of the study will be embodied in a printed bulletin shortly to be issued by the Wisconsin State Department of Public Instruction.

The accompanying chart presents a proposed form of organization for the administration of a school publicity campaign. An organization of this sort must of necessity be highly centralized and the various parts properly coordinated. The nature of the campaign committee proper, with its various subcommittees, the various kinds of active workers, together with the working relations of various parts of the entire staff may be gleaned from an examination of the chart. In small communities only as much of the organization as is indicated by the solid lines may be necessary, but in large communities the complete organization seems advisable.

One cannot read the accounts of campaigns that succeeded and of campaigns that failed without becoming impressed with the necessity for reaching all parts of the public. This may be partly accomplished by having on the campaign committee representatives of various groups such as labor, capital, real estate interests, and religious organizations. Experience indicates that special effort should be made to reach the following: (1) *Illiterates* as distinguished from literates. (2) *Those unfamiliar with the English language* as distinguished from those able to read and understand it. (3) *Women.* Women are not so much inclined to

exercise their voting privileges to ward off increased taxes, as to secure improved school conditions. They can usually be enlisted for active, enthusiastic, and effective support. (4) *Heavy taxpayers.* This group includes the bankers, real estate owners, manufacturers, etc. Endorsements from those upon whom the burdens fall heaviest will have much weight. (5) *Retired farmers.* This group is an important one to consider in smaller communities because of its notorious opposition to high taxes. (6) *Families with children of school age* as opposed to families without children or with children beyond school age. Winning the support of the latter will be much more difficult. (7) *Workers who cannot leave their work* and who must be reached either going to and from their place of employment, or at it. (8) *Young people between the ages of 18 and 20* or thereabouts. They have such interests as athletics, gymnastics, night schools, and community center activities and become staunch supporters when properly approached. (9) *School children.* The pupils can be used to influence parents. It should be remembered also that they will later become the parents. Efforts made to train them for effective support will count also in campaigns to come.

For dealing with the various groups, the best procedure is to try to win as many active supporters in each group as possible. The greater the number of intelligent and aggressive supporters that can be won, the greater will be the chances of success with the rank and file in that group. Proper timing of the campaign so as to fit in with campaigns for other purposes is often essential to secure the support of certain influential groups.

The fundamental principles for the selection of arguments and illustrations in the campaigns appear to be these:

1. Build all on the proposition that the people of the community wish to do the right thing by the children and that they will make any necessary sacrifices to this end if needs are clearly and forcibly shown.

2. Make the good of the children the paramount issue, in particular subordinating all considerations of cost to it.

3. Let the school authorities appear to be speaking for the whole people on school matters and not in the interests of any special class.

Other good rules for the selection of arguments and illustrations may be stated thus: (1) Follow the principles of good argumentation. (2) Establish confidence in the schools by showing their accomplishments. (3) Present actual facts that form the basis for present needs. (4) Bring out clearly what the new program will accomplish. (5) Secure public endorsements by leaders. (6) Treat proposed costs and the ability of the community fairly, and at the psychological moment. (7) Make a list of all possible objections and alternative propositions; have a sound and dignified rebuttal for each.

It is fully as important to prepare and circulate the publicity material properly after it is selected as it is to select good arguments and illustrations in the first place. But by far the best way to learn how to prepare the *form* of the publicity material, is to examine carefully samples of effective school publicity, with a view to imitating or reproducing them for one's own school system. Space forbids any attempt on our part, however, to present such samples here. It would be as futile for us even to try to summarize the work of preparation in a few words, as it would be to tell a person in a sentence how to be "interesting." Any treatment of this phase of the study that is to be of real service to the reader must contain abundant examples and illustrations from actual campaigns. For material of this sort the reader is accordingly referred to the forthcoming bulletin of the Wisconsin State Department of Public Instruction previously mentioned.

IV. A Composite Program

To save the busy superintendent time and effort, it has seemed advisable to map out a composite campaign with steps in chronological order, embodying the best methods of all the successful campaigns studied. In his own campaign the superintendent should strive to produce a cumulative effect with his arguments, illustrations, and citations of fact. For complete layouts of campaigns extending over several months Lincoln, Nebraska and Denver, Colorado furnish good examples.

The treatment of the composite campaign is given under three heads with the steps under each indicated in chronological order. I. The general publicity campaign (lasting from six months to

several years depending upon circumstances). II. The drive (a month or more). III. Additional work necessary to insure that the increase will actually be carried at the polls, wherever the question is submitted to popular vote.

I. The General Publicity Campaign

1. The general plans having been laid, the director and the managing committee start the discussion of school needs, preferably at a mass meeting.
2. The worth of the schools is constantly stressed; when defects must be noted, the cause is clearly shown to be mainly a lack of sufficient funds.
3. By all possible means the discussion is kept up.
4. By a survey, personal investigation on the part of citizens, and skillful comparisons, a growing conviction that something must be done, is induced.
5. The discussion is gradually directed toward a demand for a specific plan for betterment.
6. The director has the specific plan ready and when things are favorable, has it proposed by suitable sponsors.
7. All along the teaching staff have been given a sort of a "know your own schools" or a "war aims" course on the needs of the school system and the plans for the campaign. Thus as soon as the drive starts, the teachers will be ready to answer promptly and satisfactorily inquiries about school work and to serve as radiating centers for sentiment favorable to increased school support.

II. The Drive

1. The definite proposal is published, approval secured from every possible reputable organization in the city, from state department supervisors, university specialists on educational problems, school survey experts, etc.
2. The press comments favorably in editorials and runs school news in prominent places.
3. Exhibits, demonstrations, and entertainments, all duly advertised, show the work of the school.
4. Favorable interviews with all types of influential citizens appear at intervals in such fashion that the effect is cumulative.
5. News of what similar cities have done for their schools appears prominently in the papers, with all superior achievements featured.
6. Children carry to every home in town letters, circulars, and booklets explaining the proposal for increased expense, much of which material they have helped to prepare so that their parents are already greatly interested.
7. Advertisements for the school, contributed by merchants, begin and continue, increasing in size, numbers, and force.
8. At least one representative in every organization in town is listening for every note of opposition. As soon as this is reported, the rebuttal is promptly circulated where it will do the most good.
9. A poll may be taken and all favorable results published.
10. Addresses are given before every organization and at practically every general meeting, especially talks by four-minute men and children.

The talks for the latter have been prepared in school so that the parents are already interested.

11. Cartoons, special articles, and slogans appear in printed material.

12. Posters, placards, and exhibits of school work are put up in store windows, the post office, etc. Handbills and dodgers are distributed to every home.

13. Advertisements are placed on delivery wagons, vans, street cars, etc.

14. Slides are prepared for moving picture houses.

15. Specific directions for voting in favor of the issue are circulated, preferably with a picture of a ballot correctly marked and stressing the date. These are printed in papers, put on movie slides and handed around on circulars.

16. Special arrangements are made to reach factory workers at the noon hour or while they are going to and from work.

17. The Sunday before voting day every minister preaches a sermon favoring the proposition.

18. A few days before the vote, each child in school writes a letter to a relative or personal friend who is practically certain to come out to vote. If the child is working under poor school conditions, the voter is asked to vote to help this particular child to secure better schooling. If the child has a good school, the voter is asked to vote to give other children who are not so fortunate, an equal chance.

19. The day before the voting, a monster parade of school children, teachers, the board of education, and all prominent individuals and organizations supporting the measure, takes place. The object of this is to make an opponent realize that he is fighting the whole community.

20. On voting day, tags are pinned on those who have voted favorably, announcing the fact.

III. Additional Work Necessary to Insure that the Desired Increase Will Actually be Voted at the Polls

1. A complete card catalog of voters is made up through children and teachers or taken by post card, showing the way citizens will vote.
   a. Those against the measure are approached, care being taken to change their views by every legitimate means.
   b. When prominent converts are made, sufficient publicity is given the fact.
   c. Publicity is given to anything indicating that there will be a majority vote for the proposition.
2. There must be a good organization to get out the vote on election day.
   a. An automobile fleet, under a competent director to bring voters to the polls, is run by high-school boys, club women and if possible by leading business and professional men.
   b. Special effort is made to get the stay-at-home voters or those who cannot easily spare the time to go to the polls.
   c. Arrangements are made for high-school girls to stay with children and sick people so that every woman possible can cast her vote. This vote will practically always be for better school support.

THE OVERLAPPING IN THE CONTENT OF FIFTEEN SECOND READERS[1]

CLIFFORD WOODY
University of Washington

A few decades ago, practically all instruction in reading in any grade was confined to the single text selected for that particular grade. The common method of procedure was either to read a selection over and over before taking up a new one thus spending the whole year on one book, or to complete the book and spend the remainder of the year in re-reading it—often reading a book as many as five or six times. The prevailing idea underlying such practice was that constant repetition of the same subject matter made permanent the vocabulary and gave ability to interpret the printed page. With the advent of the present decade with its emphasis on social needs, economy of time, and superior methods, this practice of confining the attention to a single book is subject to much criticism. It is not denied that repetition of vocabulary is needed to guarantee mastery over it, but it is contended that greater mastery is obtained by meeting the words in many different situations than by meeting them many times in the same situation. Furthermore, reading is the key to the complete understanding of the complex social life in which we live. Adequate training in it demands a subject matter so broad and so varied as to furnish introduction to the many lines of activity in present day life. Such a training, therefore, necessitates the mastery of the reading process through extensive rather than through narrow intensive reading.

Improved methods of teaching reading involving motivation emphasize this same plea for extensive reading. Such a plea urges that constant repetition of the same subject matter deadens the interest and makes the reading process more mechanical and much less effective than repetition of the same vocabulary in different stories which may be read with the interest always at the highest possible level.

[1] All tabulations quoted in this article are taken from a master's thesis presented at the University of Washington by Ralph E. Stone. The author is indebted to Mr. Stone for the privilege of publishing them.

Simultaneously with these demands for an extensive and varied subject matter there comes upon the market an enormous supply of readers. Some are intended solely for supplementary purposes while others are intended for regular class texts, but advertised as also useful for supplementary reading. Each year thousands and thousands of dollars are spent in providing the schools with such materials. While no one questions the wisdom of an ample supply of supplementary readers, it is contended that many of these readers have been blindly and unscientifically selected and that too often readers selected for supplementary purposes contain the same stories as the basal texts.

In order to throw some light on the overlapping in the contents of second readers and to secure a few facts to aid in the selection of supplementary texts, a study involving the critical analysis of the contents of fifteen second readers was made. The titles of the books together with the names of the authors, and the dates of publication are presented in Table I.

TABLE I. TITLES OF THE FIFTEEN BOOKS SELECTED WITH THEIR AUTHORS AND DATES OF PUBLICATION

Title	Author	Date
Literary	Young and Field	1916
Aldine	Spalding and Brice	1907
Edson-Laing	Edson and Laing	1916
Progressive	Burchill, Ettinger, and Shimer	1909
Carroll and Brooks	Carroll and Brooks	1912
Baldwin and Bender	Baldwin and Bender	1911
Rational	Ward	1896
Riverside	Van Sickle and Seegmiller	1911
Elson	Elson	1913
Classic	Norvell and Haliburton	1901
Golden Treasury	Stebbins and Coolidge	1909
Beacon	Fassett	1914
Nature and Life	Smith	1910
Cyr	Cyr	1904
New Education	Demarest and Van Sickle	1901

The list of books was arbitrarily chosen but an attempt was made to select those that are widely used at the present time.

One familiar with second-grade work will recognize both the "content" and the "method" type of readers. The date of publication shows that the Rational, published in 1896, is the oldest of the books considered and that the Literary and the Edson-Laing, published in 1916, are the newest. No one of the books is intended primarily for supplementary purposes, but current practice shows that any one of them is apt to be used as a supplementary text, many classrooms having five or six sets.

The Overlapping in Types of Literature

The first phase of our problem represents an attempt to classify the contents of the different readers according to the different types of literature contained. Of course such an attempt is arbitrary but the writer is convinced that the results secured are amply reliable to show the tendency in each book.

Most of the types of literature mentioned in Table II are self-explanatory, but for the convenience of the reader, it might be well to explain what was included in such closely related types as folklore, myths, fables, and fairy stories. Folklore includes tales and legends of the Norse, English, and Scotch, as well as stories from the Grimm brothers and Joel Chandler Harris. Myths include only those selections that attempt to explain contemporary phenomena by the narrated occurrence; fables, those selections of Aesop and a few others whose purpose was to reflect the foibles and frailties of human nature; fairy stories, those in which fairies or other benevolent spirits dominate the situation.

Table II, representing the percent of each book devoted to the various types of literature, shows that the books differ a great deal in the nature of their content. Probably the greatest value of this table is to give a bird's eye view of the books, and to aid in choosing a supplementary book which will emphasize different types of literature from those already represented in the basal readers. It is surprising that a majority of books devote a preponderance of space to such closely related fields as folklore, fables, fairy stories, and myths, and a meager amount of space to stories involving boys and girls, birds, plants, pets, fowls, plays and games. It would seem at once that the primary basis for appealing to the interest of the child has been the strange and the fanciful. This fact is brought out more plainly in Table III,

TABLE II. DISTRIBUTION OF CONTENTS BY PERCENT OF PAGES IN EACH TYPE OF LITERATURE

Readers	Folklore	Fables	Boy and Girl	Elements	Rhymes	Birds	Plants	Fairies	Pets	Fowls	Plays and Games	Historical and Patriotic	Myths	Conduct	Insects	Lullabies	Miscellaneous
Literary	25.9	23.6	0.6	8.4	1.3	0.9	0.6	27.5	0.9	0.3	1.2	0.0	4.2	0.0	0.6	1.2	1.9
Aldine	26.8	5.5	2.3	6.4	0.6	6.4	11.1	0.0	0.0	26.3	0.6	0.0	1.3	3.7	4.1	3.2	2.3
Edson-Laing	38.0	4.0	5.0	0.0	8.0	6.0	0.0	22.5	0.0	5.0	3.6	0.0	6.0	2.0	0.0	0.0	0.5
Progressive	79.3	15.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.3	0.0	0.0	0.0	0.0
Carroll and Brooks	19.4	9.9	0.8	3.3	5.3	7.8	2.0	6.6	16.1	6.1	7.0	4.9	0.0	0.0	0.0	4.1	6.6
Rational	17.2	8.6	21.5	1.7	0.8	6.0	0.0	2.3	3.0	4.3	3.4	6.0	0.0	5.1	9.4	0.0	10.3
Baldwin and Bender	8.6	1.6	6.1	4.9	3.6	13.9	2.0	4.5	12.6	0.0	15.9	0.0	0.0	4.0	0.0	2.0	19.6
Riverside	38.6	6.5	3.7	1.7	1.0	3.4	0.3	16.2	1.0	5.5	3.4	7.2	4.1	3.4	1.7	0.6	1.0
Elson	44.5	7.4	6.4	2.3	0.6	3.0	3.0	18.0	1.7	0.0	0.0	3.6	4.4	0.0	0.0	0.0	4.4
Classic	50.6	2.2	0.0	3.9	0.0	6.5	2.6	6.1	2.2	5.2	1.9	10.4	3.2	0.0	0.0	0.6	3.9
Golden Treasury	15.5	7.3	3.4	2.9	0.0	10.7	1.4	0.0	10.2	4.4	0.4	0.0	2.4	0.4	2.4	2.4	24.5
Beacon	51.1	0.0	12.5	2.6	0.7	1.5	0.0	14.7	0.0	0.0	0.7	6.4	0.0	0.0	0.0	0.0	9.4
Nature and Life	0.0	2.5	11.8	3.4	1.4	13.0	14.8	0.0	7.0	4.4	0.0	2.9	1.8	0.0	3.3	0.0	33.3
Cyr	0.0	0.0	23.7	9.0	0.0	0.0	10.0	0.0	7.5	8.5	0.0	9.5	0.0	3.0	5.0	0.0	24.5
New Education	22.7	14.8	5.7	3.3	0.0	4.1	9.5	0.0	0.0	2.0	0.0	11.1	11.5	2.8	0.4	0.0	11.5
Median	25.9	6.5	5.0	3.3	0.7	6.0	2.0	4.5	1.7	4.4	0.7	3.0	2.4	0.4	0.4	0.0	6.6
Range	0–79.3	0–23.6	0–23.7	0–9.0	0–8.0	0–13.9	0–14.8	0–27.5	0–16.1	0–26.3	0–15.9	0–11.1	0–11.5	0–5.1	0–9.4	0–4.1	0–33.3
Quartile Range	14.8	3.8	5.5	1.6	0.7	3.1	4.7	8.1	3.7	2.7	1.7	3.6	2.2	1.7	1.6	1.0	8.8

which represents an attempt to generalize upon the facts given in Table II, and to group the subject matter into two main classes, the instructive and the imaginative.

TABLE III. NUMBER AND PERCENT OF PAGES DEVOTED TO INSTRUCTIVE AND IMAGINATIVE MATERIALS

Readers	Total Pages Reading Material	Pages Instructive Material	Pages Imaginative Material	Percent Instructive Material	Percent Imaginative Material
Literary	154	17.0	137.0	11.0	88.9
Aldine	108	25.5	82.5	23.6	75.9
Edson-Laing	100	16.0	84.0	16.0	84.0
Progressive	121	0.0	121.0	0.0	100.0
Carroll and Brooks	121	53.5	67.5	44.2	55.7
Rational	116	74.5	41.5	64.2	35.7
Baldwin and Bender	122	82.0	40.0	67.2	32.7
Riverside	145	25.0	120.0	17.2	82.7
Elson	147	31.5	115.5	21.4	78.5
Classic	153	49.0	104.0	32.0	67.9
Golden Treasury	102	40.9	62.0	39.2	60.7
Beacon	132	40.0	92.0	30.3	69.6
Nature and Life	135	67.5	67.5	50.0	50.0
Cyr	101	101.0	0.0	100.0	0.0
New Education	121	51.0	70.0	42.1	57.8
Median	121	40.0	82.5	32.0	67.9
Range	100-154	0-101	0-137	0-100	0-100

A glance at Table III reveals great variation in the amount of the instructive and imaginative material contained. The Progressive goes to one extreme by including no instructive material and 100 percent imaginative material, while the Cyr represents the other extreme with 100 percent instructive and no imaginative. Between these extremes one can find almost any proportion that one desires. Of all the books considered, 32 percent of the total pages represents the median amount devoted to instructive material, while 67.9 percent is devoted to imaginative.

In this connection it is interesting to indicate that on the average 15.3 percent of the subject matter in the different books—

the amounts varying from 0 to 42.0 percent—is devoted to *real nature stories;* while 51.8 percent on the average—with amounts varying from 0 to 100 percent—is devoted to folklore, fairy tales, and myths. A reader often devotes much more space to *nature myths*, the imaginative, than to *true nature stories*, the instructive. It is true that children like to read nature myths, but it is equally true that they are hungry for real facts about nature. Do these facts not indicate a lack of balance between instructive and imaginative material? Do they not suggest a definite need for the publication of a new type of reader which emphasizes the right informational material?

Duplication in the Selections Themselves

The previous section indicates great duplication in the type of material composing the contents of the different readers. One can readily understand some of this duplication if he takes into consideration the duplication in the selections themselves.

Table IV, giving the different selections occurring in at least three readers and the readers in which they occur, shows a rather surprising amount of overlapping. "Over in the Meadow" is found in the Edson-Laing, Carroll and Brooks, Rational, Baldwin and Bender, Classic, Golden Treasury, and Beacon readers. "The Four Friends," better known as "The Town Musicians," is found in the Progressive, Carroll and Brooks, Rational, Riverside, Beacon, and New Education readers. The last column in the table summarizing the number of readers containing each selection shows that one selection occurs in seven readers, one selection in six readers, one selection in five readers, six selections in four readers, and fifteen selections in three readers. A table similar to this one but not included in this article for want of space shows that 53 selections occur in two different readers. In the fifteen readers examined, 77 selections are repeated in from two to seven books.

The real significance of this overlapping is better set forth in Table V which gives the exact number of pages of duplicated material in any two of the books. To read this table always begin at the left. The figures opposite each book at the left indicate the number of pages in that book devoted to the same subject matter contained in each of the readers shown at the tops of the columns. For example, the Literary reader devotes 3.5 pages to material contained in the Aldine reader, 1.5 pages to material

TABLE IV. SELECTIONS OCCURRING IN AT LEAST THREE READERS AND THE READERS IN WHICH THEY OCCUR

Selections	Literary	Aldine	Edson Laing	Progressive	Carroll and Brooks	Rational	Baldwin and Bender	Riverside	Elson	Classic	Golden Treasury	Beacon	Nature and Life	Cyr	New Education	No. of Times Occurring in Readers
Over in the Meadow	..	..	x	..	x	x	x	..	..	x	x	x	..	..	..	7
The Four Friends	..	..	..	x	x	x	..	x	..	..	..	x	..	..	x	6
The Three Billy Goats	..	..	..	x	x	..	x	x	..	x	..	..	..	..	..	5
The Bell of Atri	..	..	..	..	..	..	..	..	x	x	x	x	..	..	..	4
The Lion and the Mouse	x	..	x	..	x	x	..	..	..	..	..	..	..	..	..	4
The Swallow	x	..	..	..	..	..	x	x	x	..	..	..	..	..	..	4
The Swing	..	..	..	..	x	..	x	..	..	x	x	..	..	..	..	4
The Wind and the Sun	x	..	..	..	x	x	..	..	..	..	..	..	..	..	x	4
The Ugly Duckling	..	..	x	x	..	..	..	..	..	..	x	..	..	..	x	4
My Shadow	..	..	..	..	..	..	x	..	x	..	x	..	..	..	..	3
Little Birdie	..	..	x	..	..	x	..	..	..	..	..	x	..	..	..	3
How to Build a Nest	..	x	x	..	..	..	..	..	x	..	..	..	..	..	..	3
The New Moon	..	..	..	..	..	x	..	..	..	..	x	x	..	..	..	3
Sleep, Baby, Sleep	x	x	..	..	..	..	..	x	..	..	..	..	..	..	..	3
Clovers	..	..	..	..	x	..	..	..	x	x	..	..	..	..	..	3
Daisies	x	..	..	..	..	..	..	..	x	x	..	..	..	..	..	3
Sour Grapes	x	..	..	..	x	..	..	..	..	..	..	..	..	..	x	3
The Dog in the Manger	..	..	..	..	..	..	..	x	x	..	..	..	..	..	x	3
The Rain	..	..	..	..	..	..	x	..	..	..	x	..	x	..	..	3
Belling the Cat	..	x	..	..	..	..	..	..	..	..	..	..	x	..	x	3
Who Stole the Bird's Nest?	..	..	..	..	..	..	x	..	..	x	..	..	x	..	..	3
Little Red Riding Hood	..	..	..	x	..	..	x	..	..	..	..	..	..	..	x	3
Whittington and His Cat	..	..	..	..	..	x	..	..	..	..	..	x	..	..	x	3
The Little Red Hen	..	..	..	..	x	x	..	..	..	..	..	..	..	..	x	3

TABLE V. PAGES DEVOTED TO DUPLICATED MATERIALS IN THE DIFFERENT READERS

Reader	Literary	Aldine	Edson-Laing	Progressive	Carroll and Brooks	Rational	Baldwin and Bender	Riverside	Elson	Classic	Golden Treasury	Beacon	Nature and Life	Cyr	New Education
Literary		3.5	1.5	3.0	7.0	3.0	0.5	1.5	1.5	6.5	0.0	0.0	0.0	0.0	5.5
Aldine	2.5		4.5	0.0	0.0	0.0	1.0	0.5	4.5	1.0	0.0	0.0	0.0	0.0	5.0
Edson-Laing	1.0	2.0		7.0	8.0	5.0	1.5	7.0	2.5	4.5	9.5	3.0	0.0	0.0	11.5
Progressive	5.5	0.0	7.0		14.5	19.0	11.5	14.5	0.0	5.5	7.0	9.0	0.0	0.0	30.0
Carroll and Brooks	7.5	0.0	6.0	10.5		17.8	7.3	10.5	2.5	13.3	2.8	11.3	0.0	0.0	12.5
Rational	5.0	0.0	7.0	6.5	15.5		5.5	4.0	0.0	3.0	3.5	16.0	0.0	0.0	14.5
Baldwin and Bender	0.5	1.0	2.5	10.5	9.5	6.5		7.5	1.0	10.5	7.0	2.5	3.5	0.0	6.0
Riverside	1.0	0.5	6.0	11.5	11.5	7.0	7.0		1.0	4.5	0.0	7.0	0.0	0.0	7.5
Elson	1.8	1.5	2.5	0.0	3.0	0.0	3.0	1.5		18.8	8.5	6.0	2.5	0.0	3.5
Classic	6.0	1.0	3.5	4.5	12.5	4.0	11.5	4.5	17.0		8.5	18.5	3.0	0.0	0.0
Golden Treasury	0.0	0.0	22.0	20.0	4.5	3.5	7.5	0.0	8.0	7.0		8.0	2.5	0.0	20.0
Beacon	0.0	0.0	15.5	9.8	17.8	29.8	2.5	9.8	4.3	17.5	7.5		0.0	0.0	18.3
Nature and Life	0.0	14.5	0.0	0.0	0.0	0.0	5.5	0.0	3.8	4.5	2.0	0.0		0.0	9.5
Cyr	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0		0.0
New Education	3.0	0.8	8.0	16.0	7.0	14.3	5.5	5.0	3.0	0.0	5.0	8.3	3.5	0.0	

NOTE: Read this table from the left when finding number of pages of duplicated matter.

contained in the Edson-Laing, etc. The presence of small numbers in these columns indicates little overlapping in any two readers, but the presence of large numbers indicates much overlapping.

Table V shows that the Cyr reader has a subject matter entirely its own and does not overlap at all with any of the other 14 readers. Such a statement cannot be made concerning the other readers. In certain cases from one-third to one-half of the contents of one reader can be duplicated without consulting more than two or three readers. To illustrate, the Progressive reader devotes 14.5 pages to subject matter occurring in the Carroll and Brooks reader and the same number of pages to subject matter contained in the Riverside reader. It includes 19 pages of material found in the Rational reader, and 30 pages of material found in the New Education reader. Again, the Golden Treasury reader devotes 22 pages to subject matter occurring in the Edson-Laing reader, 20 pages to subject matter occurring in the Progressive reader, and 20 pages to subject matter occurring in the New Education reader. Do not these facts suggest the possibility of waste in the purchase of supplementary readers?[2]

In indicating this overlapping it is not contended that children should not read and re-read the same selections for indeed they clamor for certain selections and seem never to tire of them. It is poor economy, however, both from the standpoint of the money involved and from the standpoint of introducing the child to the activities of life's complex environment to buy a supplementary reader which extensively duplicates the material already contained in the basal reader.

As stated above, children should re-read stories and they call ardently for their favorite selections, but they want exact repetitions and not abridged or modified versions in another book. Better practice, from the standpoint of modern theories of education, would be to utilize the same words in different stories involving various aspects of the social environment of the child,

[2] In pointing out the amount of overlapping, no consideration has been given to the date of publication of the different readers, nor has any attempt been made to account for the cause of the overlapping. The situation has been portrayed as it is in the hope that superintendents and teachers will be more careful in the selection of their supplementary readers.

rather than to use the same words in different versions of the same story. Such a practice would eliminate the duplication of selections.

Suggestions for Avoidance of Duplication

The facts presented in the previous sections indicate two well-defined tendencies: first, that of employing the imaginative type of literature; and, second, that of repeating the same selections in the various readers. However, with a little precaution in the selection of the basal and supplementary readers, any undesirable effects of these tendencies may be avoided.

If the basal reader is one which emphasizes the imaginative material as indicated in Table III (or through similar analysis of other books), the selection of supplementary readers emphasizing instructive material—i.e., biography, history, true stories of plants, animals, boys, and girls, etc.—would furnish variety in the types of reading materials and would harmonize with the social interpretation of education.

Reference to Tables IV and V would facilitate a choice of books without duplication of the selections themselves. To illustrate, suppose the Literary reader is the basic text and a supplementary text is to be chosen. Table III shows that 88.9 percent of the content of the Literary reader is of the imaginative type—the particular form of which can be found in Table II if desired. Obviously, if one accepts the social theory of education, the supplementary book should be of the informational type. From Table III it is seen that four readers—Rational, Baldwin and Bender, Nature and Life, and Cyr—contain more than 50 percent instructive material. Table V shows that little or no duplication exists between the Literary reader and any of these books, and consequently any one of them would make a good companion book so far as subject matter is concerned. It is evident that such comparisons become of more value as the number of readers, supplementary, companion, or basal increases.

The above tables suggest a method of analysis which should prove helpful. Until some such method is followed school people will continue to select books unscientifically and to deprive the children of reading matter of sufficient variety to introduce them to the complexities of the social life with which they are surrounded.

WHAT FIRST-GRADE CHILDREN CAN DO IN SCHOOL AS RELATED TO WHAT IS SHOWN BY MENTAL TESTS

VIRGIL E. DICKSON

Director, Bureau of Research and Guidance, Oakland, California

For the past three years, mental testing has been carried on extensively with first-grade children at Oakland. It has been discovered that a very high percentage of those who test below six years, mentally, fail to do the work which has been set up as standard for first-grade pupils. It was decided in August, 1918, to conduct an experiment by segregating some of our first-grade classes, using the results of the mental test as the chief basis for grouping. Two schools were selected where there were large numbers of first-grade pupils reported by the teachers as serious problems for education. We shall refer to these schools as School A and School B. Both schools are situated in a manufacturing, water-front district where the social conditions are among the least desirable of the city. According to a recent survey, 15 percent of all children attending School A are foreign born; 65 percent of the fathers and 61 percent of the mothers are foreign born, chiefly Austrian, Italian, and Portuguese. Of the children attending School B, 15 percent are foreign born; while 65 percent of the fathers and 57 percent of the mothers are foreign born. Here, the Portuguese element predominates.

In School A, the children of the low-first grade were segregated in three groups:

1. The borderzone group (I. Q.[1] 85 or below).

2. A group slightly better, either because of higher mental age or higher I. Q., but still below normal. This, we called the dull normal group.

3. The normal group, plus a few who tested superior.

In each of these groups a few shifts were made up or down on the basis of the school experience of the individual child. Some

[1] I. Q. refers to the "intelligence quotient" and is the ratio between mental age and actual or chronological age.

Children with I. Q. below 70. (mental age more than 30 percent below actual age) are likely to be feeble-minded, those with I. Q. between 75 and 85 ordinarily do not class with the feeble-minded, but are decidedly dull. Strictly speaking, perhaps few who test as high as 80 should be classed as borderzone cases.

of the children had been in school only a short time and others had attended one, two, or three terms. The accompanying table shows data concerning the 25 children who were placed in the lowest group. This, we shall refer to as Group I, School A.

This table includes all the pupils in the low-first grade of School A who classified in the borderzone group, mentally. Eight of these pupils tested below 70 I. Q.; ten, 70 to 79 inclusive; and seven, 80 to 87 inclusive. At the time the experiment began (August, 1918) three of the children were in the three-year mental age group, nine were in the four-year mental age group, twelve in the five-year mental age group, and one in the six-year mental age group. In chronological age, fourteen are now between seven and eight years, eight are between eight and nine years, and three between nine and ten years.

Considering the fathers of these children, eleven were born in Portugal, five in Italy, three in Austria, two in Assyria, one in Spain, one in Greece, one in England, and one in the Balkan States. Of the mothers, eight were born in Portugal, one in Honolulu (Portuguese), five in Italy, three in Austria, two in Assyria, one in Spain, three in the United States (all foreign, by nationality), one in Russia, and one in the Balkan States. From this, it will be observed that our problem involves a foreign element entirely. However all of the children spoke English rather freely at school prior to the time the segregation was made.

The social status as indicated by the occupations of the fathers is as follows: five were laborers in the ship yards; eight, laborers in the railroad shops; three, laborers on ferry boats; three, teamsters; five, day laborers; one, vegetable vender.

A strong primary teacher who was interested in this group of children was given the class. She was allowed freedom in method and equipment and told to do the best she could to educate these children. She has labored faithfully for a whole year, plus about three-fourths of the present term, with the result that only six of the 25 children can now read in an easy primer, and these do not read quite as well as the pupils in the normal class of the same school who began work this term. This statement regarding reading ability is based on an individual test for speed and accuracy in naming words in an easy primer not formerly used by the pupils. The reading of these 25 children was compared with that of twenty children who tested in the normal

GROUP I, SCHOOL A

Child's Number	Intelligence Quotient	Present Age (Nov. 1, 1919)	No. of Terms in School Prior to Present
1	2	3	4
1	78	8–7[a]	3
2	59	9–1	4
3	80	8–4	3
4	80	8–2	3
5	71	7–11	3
6	59	7–4	2
7	87	7–3	2
8	79	7–10	2
9	77	7–7	2
10	75	7–6	2
11	62	9–7	4
12	68	7–5	2
13	83	7–4	2
14	66	8–1	3
15	63	8–3	4
16	71	7–9	3
17	75	9–5	2
18	85	7–8	2
19	59	8–8	4
20	78	7–8	2
21	62	8–7	4
22	85	7–11	2
23	76	8–3	2
24	79	7–10	3
25	81	7–1	2

[a] 8–7 means eight years, seven months.

group and who had attended the first grade for only three-fourths of a term. The six pupils who have learned to read fairly well are numbers 1, 4, 5, 9, 15, and 23. Numbers 1, 4, and 5 have spent three terms each in school prior to the beginning of this term; numbers 9 and 23 have spent two terms prior to the present, and number 15 has spent four terms.

According to the teacher's judgment, these six children are the only ones of this group who now have approximately the capacity

for first-grade work. These, she believes, would not long keep pace if put in a class with normal first-grade children. They are slower than normal children and require much more drill in order to fix the work they are learning.

Space in this article will not permit detailed treatment of the data in each of the other groups. Group II, School A, was composed of 29 children who tested better than those in Group I or who showed slightly better adaptability for school work. They were considered the dull normal group. Sixteen tested in the five-year mental age group; three, in the four-year mental age group; and ten, in the six-year mental age group. The I. Q. range was from 75 to 95. This group was placed in charge of a good primary teacher, who was likewise instructed to educate these children the best she could, using great freedom in method and course of study. Six of this group had previously attended the first grade one term, eight, two terms, and two, three terms. The children have now completed two and three quarters terms of work since this experiment began, and only seven of the group have succeeded in learning to read approximately as well as the average pupil in the normal class in the high-first grade. The others can name words, but cannot do independent reading in an unfamiliar but easy first reader. Not one is considered by the teacher of this division to be able to do the standard work in a regular class of high-first grade children.

Group III of children in School A was composed of those who tested normal or above. There were 42 in this class and all but five of them passed the work of the regular first grade at the end of the term. The teacher attributed the failure of four of these to irregular attendance and of one to excessive timidity.

In School B, thirty of the pupils who tested below six years mentally and who classified in the dull normal group or below were placed in a special first-grade division. Of these, fourteen were repeating the work of the first grade, ten having been in school two terms, and four, one term. The teacher of this class was instructed to do work that was as near as possible the standard for the first grade. At the end of two terms of work under the experiment, two pupils out of the thirty were promoted into the high-first-grade regular class. Now, near the end of the third term, the teacher reports five more ready for high-first-grade work. The rest are not yet ready to be put into such a class.

Each of the three groups testing below normal has been assigned to a special teacher who has remained with that group for one year and a half, and the plan is to have the group with the same teacher for a period of two years in order that the progress of the experiment may be consistent for that time.

The superintendent of schools has appointed a committee consisting of the supervisor of primary grades, the teachers involved in these divisions under experiment, and three other primary teachers to follow this experiment carefully and report.

The committee arrived at the following tentative conclusions at the end of the first year (June 20, 1919):

1. That children who test below six years, mentally, and who also show by classroom behavior that they are below average in mental alertness should be segregated near the beginning of the first grade into a special division not to exceed 25 in each group and that the regular first-grade work should not be attempted.
2. A course of study should be made, adapted to their needs.
3. These children should move up through grade work more slowly, probably, than normal children, but not so slowly as that they will become discouraged through lack of progress.
4. The daily program of the teacher should be sufficiently flexible to permit the development each day of the problems of most vital concern to the class on that day.
5. The next problem for the committee to attack is the making of a course of study suited to the needs of such pupils.

According to the teachers of the three classes of mentally retarded, the first problems in the education of these children involve the teaching of cleanliness, hygiene, social adaptability, property rights, and the removal of fear of corporal punishment. Particularly do they emphasize the fact that these children get so many knocks and blows from their associates on the street and in the home, that they become victims of an attitude of fear and distrust which is very hard to overcome in the classroom. The customary thing is for the child to shrink or dodge whenever the teacher or any one else approaches.

An effort has been made to teach as much as possible of reading and number work, and to increase the vocabulary through conversation, dramatization, story telling, games, and handwork. The children are very slow to comprehend, slow to act, and what seems to be learned one day is lost before the next. Memory

and retention seem very poor; therefore repetition and drill must be used more extensively than with normal children.

The policy proposed by the committee mentioned above is being gradually introduced into the schools of Oakland wherever the conditions permit, and the committee has been instructed by the superintendent to proceed in the development of a course of study for the first and second years of schooling which shall be adapted to the needs of children who show by test that they are mentally retarded and who also show by brief trial that they are unable to attack, successfully, the regular work of the first grade.

Summary and Conclusions

1. Mental tests of several hundred first-grade children have shown that mental age and I. Q. are important factors in revealing a child's chances for success in his school work.

2. Most children who fail in the first grade show mental ages below six years and I. Q.'s below normal.

3. An experiment covering a year and a half leads us to approve the plan of segregating pupils in the first grade according to mental capacity.

4. The individual mental test (Stanford Revision) is trustworthy in determining such segregation in almost all cases.

5. Children who test low are very slow to learn to read. They have little initiative. What they appear to learn one day is not retained to the next. Much drill and repetition are necessary. Their reading is marked by a tendency to name words without thought of the meaning.

6. A special course of study should be given to children who show serious mental retardation. They should be led up through their school course according to their capacities and needs rather than according to formal grade standards.

7. Segregation of those who are retarded should occur during the first term or at the end of the first term in school, after the child has been "exposed" to the regular work. (Some of our committee feel that the segregation should take place immediately after the child's entrance to school.)

8. Several hundred children in Oakland are now segregated into special groups according to mental capacity and the plan is being rapidly extended.

9. A committee is now at work arranging a course of study to meet the needs of retarded children in their first two school years.

THE CAMPAIGN METHOD IN ELEMENTARY EDUCATION[1]

J. W. RICHARDSON
Research Director, Hibbing, Minnesota

The value of standardized tests for the purpose of checking up the results of classroom instruction are recognized by all of us. We also recognize that the method of using the tests has become somewhat "standardized." Briefly, supervisors and teachers value them as a basis of judgment respecting the accomplishment of a given task. If test results are high, all are satisfied. If they are low, word is passed to the teachers to emphasize this or that phase of their work a little more; and the teachers follow instructions. Let us not examine the character of this emphasis too closely. We all know what emphasizing a particular phase of a given subject means. Later the test is repeated, each teacher is raised or lowered in rank according to the findings, and the incident is closed.

NATURE OF THE CAMPAIGN METHOD

But what of that other great class of school people? I refer to the pupils themselves. What use do they make of the tests? What do they know of grade standards and the desirability of attaining to them? What would they do about it if they did know? Of course, we attempt to stimulate interest in the immediate test, and I suspect that many a teacher warns her class occasionally that another test may be given sometime and that she wants the class prepared. But however much she may exhort the class group to make a good class score, and however much we may claim that group tests are designed to "measure the efficiency of the entire school, not the individual ability of the few"; yet it is the individual who must take the tests. In the intermediate grades we have the beginnings of group consciousness, but the individual is the dominant factor. Hence, from a theoretical standpoint we should be warranted in considering individual pupils even more than we are warranted in considering individual teachers. From a practical standpoint, we have the

[1] An address delivered before the National Association of Directors of Educational Research at Cleveland, Ohio, February 26, 1920.

evidence of the interest on the part of pupils when handwriting scales are displayed in the school room.

It is to call attention to this individual phase of the work and to report upon one of its possibilities that this paper is written. And it is because of the nature of the administrative details of the plan used to secure evidence, that we have chosen our topic—"The Campaign Method in Elementary Education."

A campaign is characterized: first, by a recognized need or desire on the part of those who organize it; second, by definite time limits, particularly for the close; third, by a definitely organized effort to stimulate and sustain interest. A fourth characteristic rests upon the fact that the success or failure of a campaign is determined by individuals other than those who organize it, whether it be by voters in the political campaign, converts in the religious revival, or subscribers to the fund. Furthermore the campaign has a peculiar value in awakening the individual to the duties and privileges of cooperation whether civic, social, moral, or intellectual.

This report will be concerned with three campaigns carried on in alternate years for the purpose of improving the work in the four fundamental processes in whole numbers, although we have some evidence to warrant the belief that the plan is just as feasible in the formal or mechanical phases of other elementary school work.

The First Campaign

When we took charge of educational research in the Hibbing, Minnesota schools in the fall of 1915, one of our first acts was, of course, to give the Courtis Standard Research Tests in Arithmetic, Series B. We found that scores were considerably below standard. Accordingly the teachers were called together and asked to "emphasize" the work in the four fundamental processes with whole numbers. Then we turned our personal attention to other subjects.

In January of that school year, it was decided to introduce the departmental plan at one of our buildings. Ten intermediate-grade classes each having sixty-minute periods for recitation and supervised study were in the departmental group. With the introduction of this plan, another form of the Series B Tests was given. Nine school weeks with "emphasized" work in arith-

metic had elapsed since the first test. The improvement, however, in the use of whole numbers was slight, except that some increase in accuracy had been made in several cases. It was decided that "something must be done." In conference with the two teachers, each of whom had five arithmetic classes daily, it was decided to start work first with the pupils who had made the lowest scores.

On large sheets were placed the names of all the pupils and following each name were the scores made by that pupil in the November and January tests, with blank spaces for another test which was to follow. The date of this third test had not been set. Hence, one of the four chief characteristics of a campaign was lacking. The blank spaces after each name, however, tended to make a personal appeal for preparedness. These name sheets, together with charts showing class medians, were placed on the walls in the arithmetic rooms.

Next, the teachers listed, for their own information, the names of the pupils who had made the lowest scores. These pupils were told that they must do outside work. Thus, in the beginning we overlooked another characteristic of the true campaign. Instead of securing action by stimulating interest we were going to force action.

Then the drive started. The pupils on the low score list were furnished mimeographed sheets of examples to work at home. They also put in extra time before and after school. Pupils who were not required to do this extra work, when they saw other pupils doing so much, became concerned about their own standing. The result was that they, too, asked for drill sheets to take home. Then they commenced coming a little earlier in the morning, and soon they were working as industriously as the others. From this time on the work was largely voluntary on the part of the slower pupils as well. They timed themselves and each other, and estimated, as closely as possible, the improvement due to practice.

The pupils were shown how to make individual charts and these appealed to them even more than did the numbers representing their scores. The standard was represented by a red line and the pupil's performance by a green one. "Raise the green above the red" became the slogan.

The pupils had become so intensely interested in this phase of the work, and the amount of class time given to it was so limited

(five to eight minutes daily), that we saw no reason for stopping it. But at the end of eleven weeks the test was repeated.

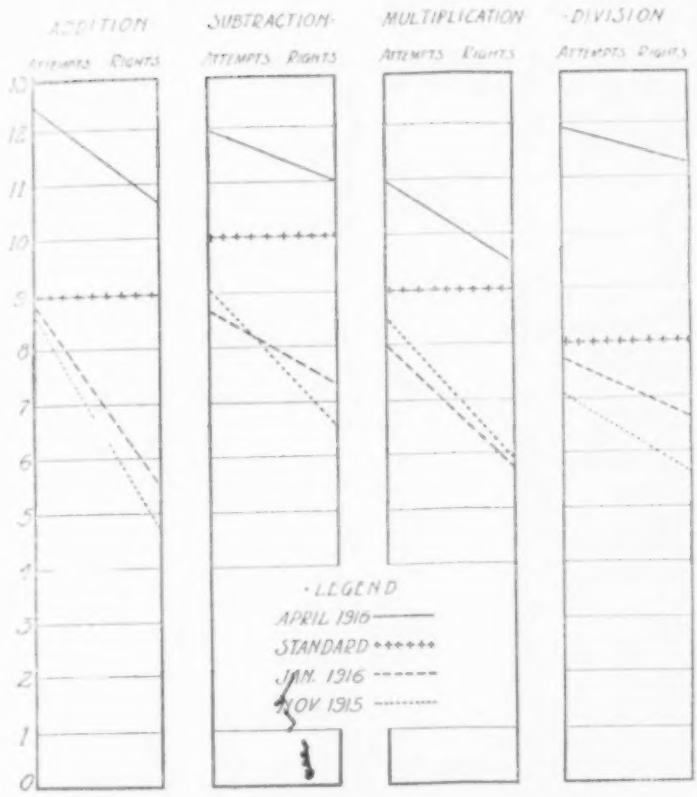
In this discussion we shall not give detailed figures. Briefly, however, it may be stated that the improvement in rate of working represented a gain of at least a year and a half and in several cases two years or more, if we consider a gain of two examples as a normal year's progress. But a still greater gain was made in accuracy. In January, at the beginning of the campaign, out of four chances (one in each fundamental operation) for each of ten classes, or out of forty chances in all, only two class medians for accuracy reached or exceeded 90 percent, while fifteen were below 70 percent. But in April at the end of the campaign, sixteen or 40 percent of the cases were above 90 percent in accuracy and none were below 70 percent. The amount of improvement in the fifth and sixth grades for both rate and accuracy is graphically shown in Figures 1 and 2.

The campaign results fell short of the original plans in one particular. We expected to boost the makers of low scores up to the class median, or up to the standard for the grade if the class median was too low, and let the rest of the pupils alone. In that way we would secure closely graded, homogeneous groups in respect to the four fundamental processes in whole numbers. But class variation increased rather than decreased. That is, there was a greater difference in the number of examples worked by the highest and the lowest pupils at the end of the intensive period than there had been at the beginning. The point is that while the low score makers were advanced the original high score makers refused to "stay put." Truly, it has been said: "The brightest pupils are the most retarded."

The Second Campaign

The next school year the tests were not used. But in the fall of 1917 we organized our second campaign. This time all four of the characteristics of a true campaign were observed. First, in order that the regular work might proceed under the most favorable conditions, we felt the desirability of bringing the work of the intermediate grades to a relatively high degree of excellence as early in the year as possible. Second, definite time limits were set. The campaign was to start just one month after school

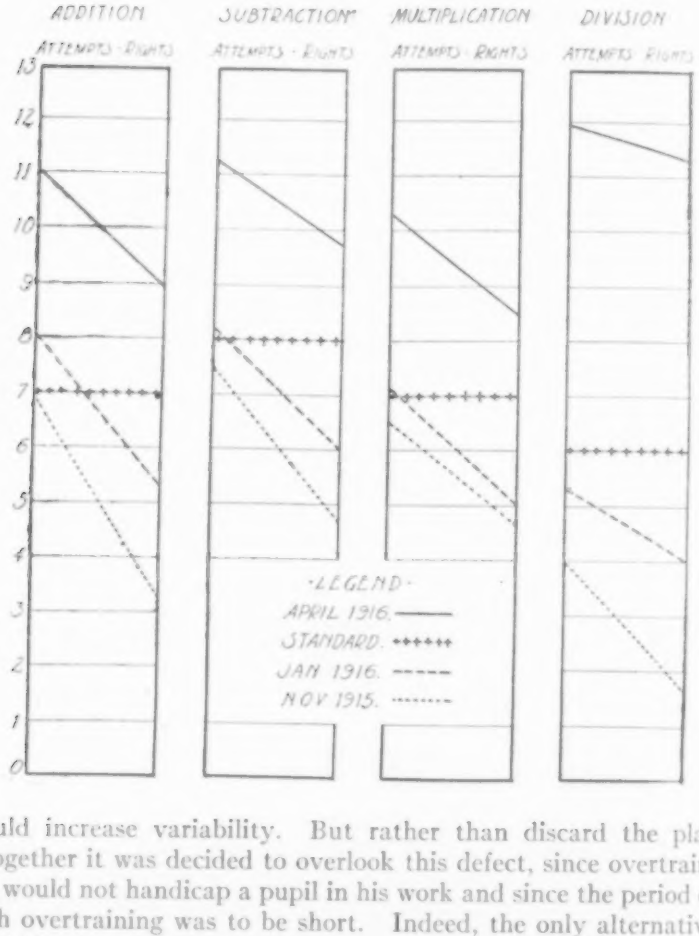
FIGURE 1. DEVELOPMENT OF FOUR SIXTH GRADE CLASSES
NINE SCHOOL WEEKS BETWEEN NOV. AND JAN. TESTS
ELEVEN SCHOOL WEEKS BETWEEN JAN. AND APR. TESTS



opened and was to close in exactly six weeks, finishing, therefore, just before Thanksgiving. Third, we made an organized effort to stimulate and sustain interest throughout the period; and fourth, of course, success or failure would depend upon pupil performance, i.e., would depend upon individuals other than the organizers.

In planning the second campaign so many factors were found to need study that we arbitrarily selected some and discarded others. For example, we knew that wide variation existed in each class and we knew that the plan followed in the first study

FIGURE 2. DEVELOPMENT OF FIVE FIFTH GRADE CLASSES.
·NINE SCHOOL WEEKS BETWEEN NOV. AND JAN. TESTS·
·ELEVEN SCHOOL WEEKS BETWEEN JAN. AND APR. TESTS·



would increase variability. But rather than discard the plan altogether it was decided to overlook this defect, since overtraining would not handicap a pupil in his work and since the period of such overtraining was to be short. Indeed, the only alternative would have been to discard the plan entirely, since the success of any campaign depends upon "boosters" among the people campaigned. Witness the preponderance of Democrats at a democratic rally or of Christian people at a religious revival. We determined, however, to obtain some check upon this phase of

the problem by observing the amount of voluntary work done by each pupil.

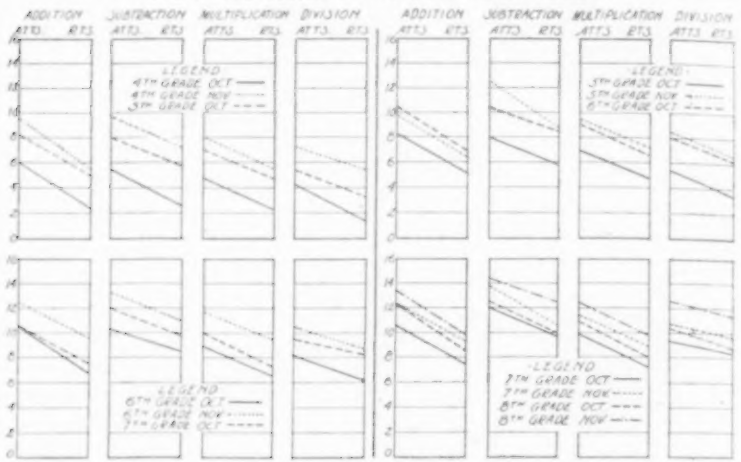
This time the work extended from the fourth to the eighth grades inclusive and in all schools where these grades were found. This included some fifteen schools, several of them one- or two-teacher schools. The number of pupils in each grade ranged in round numbers from 250 in the fourth grade to 150 in the eighth.

The amount of class time given to drill work was limited to ten minutes daily except in the case of the fourth grade in long division. The same general plans were followed as before. Lists of pupils' scores and class charts were displayed. Pupils were furnished all the examples they wanted for outside work, except that each teacher was asked to use her judgment in limiting the number given to individuals whom she considered physically unfit for outside work. But this time pupils were allowed to record the number of examples which they turned in for correction. For the most part these examples were never corrected. This matter will be touched upon again in connection with the third campaign.

The results of the second campaign were clear cut and decisive. The first test (the one given at the beginning of the second campaign) revealed the fact that the pupils were well graded in respect to ability in fundamental operations with integers, since there was, in general, a difference of one to two examples between each two successive grades and since in only one case did one grade median equal that of the next higher grade in any of the four operations. Furthermore, each grade median in rate with the exception of those of the fourth grade in subtraction and multiplication, was within one example of the Courtis standard for the grade. The second test given at the end of the six-week period, showed that, with but one exception, each grade median was above that which the next higher grade had earned at the beginning of the period. The one exception was the fifth grade in addition. At the beginning of the campaign the median of all median accuracies for all five grades in the four operations was 70 percent; at the end it was 78 percent with only two cases below 70 percent, these being in the sixties. At the end of the period, each grade, judged by the performance of its median pupil, was better prepared to deal with integers than the grade above had

been at the beginning of the campaign. These facts are shown in Figure 3. In terms of their own performance, all grades combined in all operations combined, improved the rate of work one-fourth and worked one-third more examples correctly.

FIGURE 3 – RESULTS OF SECOND CAMPAIGN - SHOWING HOW EACH GRADE SURPASSES THE INITIAL RECORD OF THE NEXT HIGHER GRADE.



As stated before, the pupils kept records of the number of examples worked outside of class. The pupils of each grade were grouped on this basis. For each process, Group I was made up of pupils who reported that they worked less than one hundred examples; Group II reported 100 to 199 examples; Group III, 200 to 299 examples; and Group IV, 300 or more examples. The group medians in each grade and process were found for the tests given at the beginning and at the end of the six weeks' campaign. There were four groups for each grade, and each group had four medians, one in each process, except the fourth grade in division. Thus there were nineteen medians for Group I—three for the fourth grade, and four each for the succeeding grades. Similarly there were nineteen medians for Group II, nineteen for Group III, and nineteen for Group IV. Taking the median of these medians for each group, figures are obtained that represent respectively the performances of pupils who worked less than 100 examples, 100 to 199 examples, 200 to 299 examples, and 300 or more examples.

These figures reveal a high positive correlation between the number of examples worked and progress. In rate of working, the first group of pupils (who worked less than one hundred examples) gained less than two examples; the second group gained a little more than two, the third three, and the fourth four and one-half, or two years' growth in the six weeks. In the number of examples worked correctly, the gains were 1.2 for Group I; 2.7 for Group II; 3.5 for Group III; and 4.5 for Group IV.

Compared with Group I, the most industrious group not only made more than twice the gain in rate, and almost four times the gain in number worked correctly, but these gains were also made on the basis of a higher initial score. In the beginning there was a difference of approximately two examples in rate and one example in number correct between Groups I and IV.

The Third Campaign

We intended to use this information in the third campaign to be conducted in the fall of 1919. We were going to give each pupil, immediately following the initial test, a personal folder containing the number of examples in each process which he should work to bring his score up to the grade standard. This could be so handled that it would prove an additional incentive to work.

But as the time for the third campaign approached, a large part of the writer's time was taken up somewhat unexpectedly with administrative details in connection with our outlying schools. It was found impossible to give even the usual amount of attention to the campaign just at that time. Hence, it was decided to proceed with a very simple form of organization. The teachers gave the tests. Clerks scored the papers and made out individual score lists which were displayed in the arithmetic rooms. Announcement was made to the teachers that the campaign would close in six weeks, and they were asked to encourage the pupils to practice for the purpose of raising their scores. They were also asked to limit class-time drill to ten minutes daily. They were furnished with self-keyed drill cards for the outside work; and by a judicious distribution of these, overtraining was partially eliminated. Clerical help was provided to correct part of the papers. The amount of work turned in had made correction of papers impossible in the second campaign, but this time with

the help of self-keyed examples and some clerks, it was done. After correction, the papers were returned to the pupils who recorded the number worked correctly.

The initial and final tests were given to grades IV to VIII but the campaign was confined to the fourth, fifth, and sixth grades. The seventh and eighth grades served as controls.

The value of the plan, even when thus hastily and imperfectly organized, was apparent when the second test papers were scored. In the fourth, fifth, and sixth grades the median of all the gains made by all grades in all processes was 1.7 examples both in number worked and in number worked correctly. This indicates a gain in percent of accuracy. In the seventh and eighth grades, on the other hand, the gain was one-half of one example both in number worked and in number worked correctly. Furthermore, at the beginning of the campaign in the fourth, fifth, and sixth grades, no grade median in any process was equal to the Courtis standard for the grade except the fifth-grade medians in subtraction and division which exceeded it by one-tenth and two-tenths of one example respectively. But at the end of the campaign each grade median equaled or exceeded the standard in each process except the fourth grade in multiplication which missed it by one-tenth of one example. We are unable to report evidence of any value in having the pupils count only the examples they had correct, rather than having them count all examples worked. Using the same method of computing gain as that used for the second campaign, the gain in accuracy was only one-half as great as in the second campaign. This feature of the work should not be discarded, however, without further study because in the second campaign all pupils had been campaigned indiscriminately, while in the third campaign the subject was not emphasized as much in the better classes as in the poorer ones.

Permanency of Acquisition

The second campaign was conducted between October 3, 1917 and November 15, 1917. In June of that year a third form of the Courtis test was given. The object was to find out how much of the gain made during the second campaign had proved to be permanent. Only the scores of pupils who took all the tests were considered. Approximately the same results were obtained as at the end of the campaign in November. Slight additional

gains were made in the lower grades, while small losses were sustained in the upper grades. It will be recalled that in the six weeks of the second campaign pupils had in general improved one-fourth in rate of work and one-third in accuracy. The June test showed that the effect had lasted seven months.

The third campaign, which was conducted in the fall of 1919, afforded an opportunity to estimate the permanency of these gains over a still longer period. Two years had elapsed. The fourth, fifth, and sixth grades had now become the sixth, seventh, and eighth respectively. The first test in the third campaign showed in these grades the effects of the training two years before. In but three cases were these grades below the present grade standards as much as one example, even within a month after the beginning of the school year. In another month, without any attention to the subject in the seventh and eighth grades, except the giving of the tests, no grade median in any process was found to be below the present standard as much as one example.

Summary and Conclusions

To summarize, the first campaign with ten classes showed us the possibilities of the plan, but eleven weeks' work resulted in overtraining for most of the pupils. The second campaign, carefully organized and administered, caused each grade to attain a performance in excess of that of the next higher grade at the beginning of the campaign. This campaign also proved the positive correlation between amount of work and progress. The third campaign, hastily organized, and administered wholly by teachers, resulted in a gain of almost two examples. And from the first test of this campaign was obtained some evidence of the permanent value of the previous one.

In discussing the characteristics of a campaign, we mentioned its value as a means of awakening individuals to the duties and privileges of cooperation. It is impossible to gauge this value exactly any more than we can calculate the degree of national awakening due to the Liberty Loan campaigns or the broadening of human sympathies due to the great war-relief drives. Yet it is certain that these school campaigns were valuable in fostering the spirit of cooperation. Pupils sought aid of each other, and in turn, offered help; parents were asked by their children to time and examine the home work; and the schools themselves came in

for their share of parental attention. If you doubt this, start a campaign in your school in arithmetic or geography or spelling or oral English. But allow me to suggest the inadvisability of starting more than one at a time, or of following one too closely by another.

Any one who has ever seen pupils of the intermediate and grammar grades gathered around a chart dealing with any aspect of their school work ranging from school attendance to classroom performance will not deny the possibilities of a further use of such charts. Children, like the rest of us, want to know just what their task is, and particularly what the limits are. Moreover, they work best when they know that their own performance rather than some teacher's judgment will decide the success or failure of their efforts. We have tended too much, in school work, to emphasize the lower limit—the passing mark—and to refuse to acknowledge the possibility of any pupils reaching the upper limit—100 percent. Children refuse to accept this view of their own limitations, but what can they do about it? Even when the teacher knows exactly what she wants, the pupils seldom know. Their task is to keep busy and avoid making trouble. Their problem is analogous to what ours would be if we were sent to a machine shop or a prescription counter or a lawyer's office and told to keep busy, to learn as much as possible about mechanics or drugs or the law, and, above all, not to disturb anyone else. But how different the pupil's problem becomes when he knows exactly what he can do now, and what he is challenged to be able to do at a given time; and when he knows that the success or failure of his efforts will be determined by his own performance at that time.

Allow me to say in conclusion that although we do not yet know the limitations of the plan, we do know that we can cause a grade to make a year's normal progress in the use of whole numbers in six weeks. And although we do not hail it as a panacea for all the ills of elementary education, we are firmly convinced of the value of the campaign plan in all aspects of school work which depend upon the memorizing process or the acquisition of skill, or in any field where definite standards can be established and reliable measures of attainment are available.

EDUCATIONAL MEASUREMENTS IN A SMALL SCHOOL SYSTEM

FRANK H. KOOS
Superintendent of Schools, Anoka, Minnesota

Really to know how the results of teaching fundamental subjects in their school system compared with the work of other organizations over the country became the desire of the teachers of Anoka city schools. They had been told that the schools were the best in the land by patrons and by school board members. The state inspectors who spent from half a day to a day each year with them said that the instruction was good. But these snap judgments hardly sufficed to give any satisfaction in the face of present day educational science.

The desire to know caused these teachers to measure their work by standards which were being used in other school systems. They made a study of many tests and school surveys in order to secure means for testing and data for comparison. They used as texts *Educational Tests and Measurements* by Monroe, DeVoss, and Kelly, and *The Scientific Measurement of Classroom Products* by Chapman and Rush. The instructions and directions for all tests were studied very carefully. It was finally decided that the school work should be measured by the following scales and tests:

1. Ayres Handwriting Scale, "Gettysburg Edition"
2. Ayres Spelling Scale
3. Nassau County Supplement to the Hillegas Composition Scale
4. Thorndike Scale Alpha 2 for Measuring the Understanding of Sentences
5. Kansas Silent Reading Tests
6. Courtis Standard Research Tests in Arithmetic
7. Stone Reasoning Tests
8. Trabue Completion-Test Language Scales.

In order to make the giving and grading of the tests as uniform as possible and in order that the burden should not be too great upon any group of individuals, a number of teachers' meetings were conducted. These served to familiarize all the teachers with the giving and scoring of tests and papers. In every case the

teachers' results were checked by one to four competent persons. The board of education thought the work of so much worth that it secured extra clerical help for checking and tabulating.

The work began in March, 1919, and ended early in May. It consumed a great deal of time and energy but the satisfaction of knowing where we were weak and where we were strong more than repaid us for all our pains.

Because spelling tests were the easiest to correct, we began with that branch of the work. We found that the Springfield, Butte, and Elyria school surveys used those word columns which demanded a spelling ability of 73. Accordingly, words were taken from those lists. Table I shows the median scores made by grades II to VIII. It also gives the scores made by other school systems. We find by comparison that Anoka stands a little better in almost every grade (except the third) than three other systems and worse than two. That is, the teaching of spelling may be characterized as average except in the third-grade classes. In these classes, the results were poor.

The fact that the third-grade classes fell so far below the rest of the system demanded explanation and caused investigation. From the age-grade table we found that the average age for the fourth-grade classes at the beginning of the year was nine years and five months. The average of the third-grade classes was nine years and eight months. This indicated retardation. Upon investigation with the opposites and the Binet-Simon tests, eleven subnormal individuals were found in this grade. This may have been the cause of the poor spelling. However, other causes were found to be the lack of suitable lists of words and poor teaching methods. There was too little individual attention, too much repetition of easy words, too little repetition of difficult words, too much oral spelling, and not enough sentence dictation.

The tests brought to us better methods of teaching spelling. The sixteenth and eighteenth yearbooks of the National Society for the Study of Education, all other available material upon the teaching of spelling, and word lists were brought into use. Numerous tests and contests were undertaken. During September, 1919, tests were given again. Another sampling will be taken in March, 1920, in order to measure improvement.

TABLE I. AYRES SPELLING SCALE—MEDIAN SCORE

Grade	Anoka			Springfield	Butte	Columbia	Elyria	Portland	Ayres
	Franklin School	Washington School	Mean						
II...	77	89	83	70	86	69	48	90	73
III...	59	37	48	65	82	66	64	79	73
IV...	65	81	73	70	79	70	67	88	73
V...	74	82	78	72	85	68	69	87	73
VI...	80	75	78	68	75	74	75	87	73
VII...	77	81	79	73	76	69	72	82	73
VIII...	74	.....	74	75	89	67	67	86	73

Because we could secure results for comparison from the various sources and because of the ease of scoring, we adopted the Kansas Silent Reading Tests to measure our results in silent reading. Table II gives the median results for the two Anoka grade buildings. With the exceptions of one third grade, one seventh and one eighth grade, the scores stand well above the average. This third-grade score is low because the whole group had been divided into divisions upon an intelligence basis. The section which made the low marks was the slowest moving division. The low record made by the Washington seventh grade may be accounted for by the fact that the regular teacher was called home on account of sickness. She and her substitute were going over the week's plans while the youngsters were taking the tests. We could in no way account for the work done on this test by one eighth grade. In order to find further proof for poor reading ability in the two grades mentioned, we administered the Thorndike Scale Alpha 2 for Measuring the Understanding of Sentences. The score made by the children was that standard of "8" given by Thorndike for the second year in high school. However, as we know nothing about the correlation of these tests, one with another we cannot say anything concerning the reading ability of these children other than that they stood low according to the Kansas standard and high according to the Thorndike standard. We did derive a great deal of satisfaction from the fact that the youngsters could read well according to Thorndike.

TABLE II. KANSAS SILENT READING TEST—MEDIAN SCORES

Grade	Anoka			Kansas	Iowa	Far West	Southern	Author
	Franklin School	Washington School	Mean					
II	5.2	4.75	5.0	.....	.....	.....	.....	.....
III	2.3 9.5	10.6	7.8	4.9	6.2	6.1	4.7	5.3
IV	14.6	13.0	13.8	9.0	9.5	10.6	8.4	9.5
V	14.6	16.3	15.4	13.4	14.6	14.4	12.3	13.2
VI	13.0	14.9	14.0	13.7	14.8	15.0	11.8	13.9
VII	22.3	13.5	17.9	16.1	17.7	18.0	15.4	16.2
VIII	19.7 15.8	.....	17.7	20.1	20.6	20.6	19.2	19.2

Even though our scores were high throughout, our teachers were not satisfied with their teaching of reading. They are endeavoring to learn the most scientific methods of teaching this subject. Silent reading now gets more emphasis. This fact is evidenced by the use made of lists of supplementary readers, the requisitions for new material, the increased use of the various libraries, the call for much more new library material (both books and magazines), and the lists of outside reading posted and used in each grade room.

Table III gives the median results for the Stone Reasoning Tests. These tests were corrected and graded by three people. In giving the tests, the teachers were especially careful to follow directions. The results obtained were the most gratifying of all. The scores no doubt demonstrate the reading ability of the pupils as well as their reasoning power. The median accuracy scores are higher than the median reasoning scores of all the other cities except those of Salt Lake City. The scores may indicate that in the study of arithmetic, too much emphasis has been placed upon this type of work. The results for the Courtis Standard Research Tests in Arithmetic, Series B, also suggest that perhaps too much time and emphasis are devoted to reasoning problems and not enough to habit formation in the four fundamentals. While the speed and accuracy scores are not the lowest within the lists of comparisons, they fall far short of the author's standards and the better work of some of the schools listed. Table IV shows the

TABLE III. STONE REASONING TESTS—MEDIAN SCORES

Grade	Reasoning Scores							Accuracy Scores		
	Anoka			Butte	Salt Lake City	Elgin	Janesville	Anoka		
	Franklin	Washington	Mean					Franklin	Washington	Mean
V	4.2	4.0	4.1	2.2	3.7	2.9	2.4	3.2	4.0	3.6
VI	6.4	5.0	5.7	3.9	6.4	5.0	3.4	5.5	4.0	4.7
VII	6.8	10.8	8.8	5.8	8.6	6.0	5.8	6.8	9.9	7.9
VIII	9.7	.....	9.7	7.7	10.5	8.0	6.3	9.4	.....	9.4

median number of examples correct at Anoka and a comparison of these figures with those obtained elsewhere and with Mr. Courtis' standards. The comparisons for the number of examples attempted and for the percent of accuracy indicate about the same relative position for Anoka. The results informed us that much time, energy, information, and teaching ability will be necessary to bring our schools up to standards set by other schools and by the needs of ordinary life. As soon as the scores were reported the teachers began to call for all types of information upon teaching and drilling fundamental facts in number work. Four sets of Studebaker Economy Practice Exercises in Arithmetic[1] were purchased and are now used. The sixteenth and eighteenth yearbooks of the National Society for the Study of Education were again brought into service. Better teaching of the fundamentals will result.

The product of a handwriting test was judged both for speed and quality by means of the Gettysburg Edition of the Ayres Handwriting Scale. The work was done by two teachers, the writing supervisor, and the teacher of commercial subjects. The average of their scores was taken. As the median marks indicate, the speed of the pupils is high and the quality is low. The only remedy we can find for this condition is not less attention to speed but more attention to letter and word formation. By a study of the quality (Table V) we notice that there is no marked improvement in ability to write more legibly until we reach the seventh grade. This grade as well as the eighth suddenly approach the standard in quality. The cause for such poor penmanship in the

[1] These are published by Scott, Foresman and Company.

lower grades and such sudden improvement in the upper grades may be due to the writing method used. We feel that we wish to develop rapid writers rather than slower drawers. We have not yet found a method whereby we can keep our speed and at the same time teach better quality.

TABLE IV. COURTIS STANDARD RESEARCH TESTS IN ARITHMETIC, SERIES B—RIGHTS

Grade	Anoka			Detroit	Boston	Indiana	Salt Lake City	Butte	Columbia	Kansas	Author
	Franklin	Washington	Mean								
Addition											
IV	2	2	2	.....	.....	.....	.....	.....	.....	.....	6
V	4.2	3	3.6	5.7	6.6	3.6	4.1	2.9	2.4	3	8
VI	4.8	4.9	4.8	7	8.3	4.4	6.4	3.4	2.9	3.9	10
VII	5	4	4.5	7.5	9.0	4.9	6.9	3.8	3.9	4.8	11
VIII	6	.....	6	9.4	10.4	5.8	8.5	5.3	4.0	5.4	12
Subtraction											
IV	3.5	4	3.7	.....	.....	.....	.....	.....	.....	.....	7
V	7	6	6.5	7.9	7.7	5	5.2	5.5	4	4	9
VI	6.4	6	6.2	8.6	9.5	6.5	7.8	5.8	4.8	5.9	11
VII	8	7	7.5	9.9	10.3	7.9	8.8	7.1	5.6	7.2	12
VIII	12	.....	12	12.5	11.8	8.9	9.8	9.8	5.4	7.7	13
Multiplication											
IV	4	4	4	.....	.....	.....	.....	.....	.....	.....	6
V	6.3	4	5.2	5.8	5.6	3.9	4.3	4.1	3.4	3.1	8
VI	5.4	5	5.2	7.2	7.2	5.1	5.3	5.0	4.4	4.7	9
VII	5.8	7	6.4	7.8	8.2	5.9	7.1	6.5	5.5	5.9	10
VIII	8.4	.....	8.4	9.8	9.3	7.3	8.3	8.1	4.6	8.3	11
Division											
IV	4	2	3	.....	.....	.....	.....	.....	.....	.....	.....
V	5	3	4	4.6	4.9	2.6	3.0	3.6	2.5	2	6
VI	4.2	3	3.6	7.3	7.4	4.8	5.5	4.3	3.9	3.5	8
VII	4.6	6	5.3	9	8.8	6.7	7.7	7.2	5.5	5.3	10
VIII	8.8	.....	8.8	11.7	11.0	9.1	9.5	10.2	5.5	8.3	11

Language and composition ability were tested and scored by the Trabue Language Scales, Series B and C, and by the Nassau County Supplement to the Hillegas Scale. By the median scores in Tables VII and VIII it may be seen that the pupils stand high in language and composition writing abilities. The

compositions were scored by four judges. Two of these were teachers, the third was the superintendent, while the fourth was engaged from outside the system. The judgments were independent. The average of the four opinions was taken as the final score. While some differences in scores were found, there was a striking similarity in the four sets of marks.

TABLE V. AYRES HANDWRITING SCALE "GETTYSBURG EDITION"—MEDIAN QUALITY SCORES

Grade	Anoka			Cleveland	Kansas	Illinois	Iowa	Ayres	Gary
	Franklin	Washington	Mean						
II	30	33	31.5	.....	44	39.7	35.7	38	.....
III	30 33	30	31	.....	47	42	39.8	42	30
IV	30	33	31.5	.....	50	45.8	44.5	46	31
V	37	33	35	45	55	50.5	49.1	50	34
VI	37	47	42	48	59	54.5	52.3	54	36
VII	63	50	56.5	50	64	58.9	57	58	39
VIII	50 67	.....	58.5	55	70	62.8	61	62	42
IX	.....	.....	85	.....	.....	.....	.....	.....	43

TABLE VI. AYRES HANDWRITING SCALE—"GETTYSBURG EDITION" MEDIAN SPEED SCORES

Grade	Anoka			Cleveland	Kansas	Illinois	Iowa	Ayres	Gary
	Franklin	Washington	Mean						
II	30	35	32.5	.....	32	30.6	39.2	32	.....
III	32 42	42	38.5	.....	35	43.8	49.2	44	18
IV	55	88	71.5	.....	51	51.2	61.9	56	45
V	81	84	82.5	60	61	59.1	65.5	64	58
VI	72	84	78	70	67	62.8	72.6	70	70
VII	82	86	84	78	71	67.9	75	76	80
VIII	84 92	.....	88	80	73	73	76.5	80	90
IX	.....	.....	86	.....	.....	.....	.....	.....	100

TABLE VII. TRABUE LANGUAGE SCALES, SERIES B AND C

Grade	Author's Standards	Anoka		
		25-percentile	Median	75-percentile
II	6	10	12	16
III	12	11	13	16
IV	16	14	18	21
V	18	18	22	25
VI	22	21	25	28
VII	24.6	23	26	28
VIII	26.6	23	26	29
IX	28.4	27	28	31

TABLE VIII. NASSAU COUNTY SUPPLEMENT TO THE HILLEGAS SCALE—MEDIAN SCORES

Grade	Anoka			Salt Lake City	Butte	Author
	Washington	Franklin	Mean			
III	2.5	3.2				
IV	3.5	3.5	3.5	2.9	2.34	3.5
V	3.1	3.8	3.45	3.1	2.87	4.0
VI	4.1	4.8	4.45	3.8	3.40	4.5
VII	4.9	5.5	5.2	4.4	3.75	5.0
VIII		6.1	6.1	5.4	4.11	5.5

The measurement work above described has been very beneficial to the teachers, supervisors, principals, and superintendent. This is proved by the fact that some of the teachers have asked that during the present year two measurements be taken. The first samples were taken in October. The second will be taken in April. We shall then be able to see how much improvement has been made in the various subjects and grades.

This study of the tests has been a great help to the school system. It has acquainted the teachers with the measurement movement in education. It has helped them to discover individual weaknesses in various pupils. It has made the faculty more cognizant of individual differences. They are beginning to learn best methods for teaching fundamentals; methods based not so much upon *a priori* reasoning as upon facts, results, and investigations.

There has been much professional growth on the part of the faculty. As a manifestation of their interest sixteen of the force are taking a University of Minnesota extension course in mental diagnosis. This knowledge should bring further improvements in their teaching.

References

Cubberley, E. P. *School organization and administration.* Yonkers, N. Y.: World Book Co., 1916.

Educational survey of Elyria, Ohio. (United States Bureau of Education. Bulletin No. 15, 1918.)

Monroe, W. S., De Voss, J. C., and Kelly, F. J. *Educational tests and measurements.* New York: Houghton Mifflin Co., 1917.

Public schools of Columbia, South Carolina. (United States Bureau of Education. Bulletin No. 28, 1918.)

Public schools of Springfield, Illinois. Educational section of the Springfield survey conducted under direction of Leonard P. Ayres. New York: Division of Education, Russell Sage Foundation, 1914.

Strayer, George D. *Some problems in city school administration.* Yonkers, N. Y.: World Book Co., 1916.

SUBSTITUTING SMALL NUMBERS FOR LARGE ONES IN THE COMPUTATION OF COEFFICIENTS OF CORRELATION

LEONARD P. AYRES
Russell Sage Foundation

Great economies of time and labor in the computation of coefficients of correlation may be effected through reducing the magnitude of the numbers used, without changing the trends of the two series for which the coefficient is to be found. Taking two very short series for illustrative purposes, we may suppose that we desire to find the correlation between the following:

S	R
22	108
23	110
25	110
26	113
27	114

In computing the correlation we are ascertaining the amount of agreement between the trends of the two series. These trends are determined, not by the size of the numbers in each series, but by the amounts by which the larger numbers in each series are in excess over the smaller ones. Because of this, the trend of the first, or subject series will be unchanged if we subtract 21 from each item in it and the same will hold true for the second, or relative, series if we subtract 107 from each item. The two series now become the following:

S	R
1	1
2	3
4	3
5	6
6	7

The correlation computed from the smaller numbers is just the same as that from the larger ones and the work is much easier. The rule to follow in most cases where it is desired to find the coefficient of correlation between two listed series of paired values

is to begin by subtracting from each number of each series an amount that is smaller by one than the smallest number in the series.

In most tables of distribution, the values are stated by classes as in the following instance:

Annual Salary in Dollars	Years of Service				Total
	1 to 1.9	2 to 2.9	3 to 3.9	4 to 4.9	
600 to 699	4	1	1	.....	6
700 to 799	1	2	5	2	10
800 to 899	1	7	10	4	22
900 to 999	1	4	7	10	22
Total	7	14	23	16	60

In this case the scale of salaries runs by steps of $100 each and that of years by steps of 1 year each. The correlation should be computed from a table in which these steps have been substituted for the original values. The table will be as follows:

S	R—Years				
Salaries	1	2	3	4	Total
1	4	1	1	.....	6
2	1	2	5	2	10
3	1	7	10	4	22
4	1	4	7	10	22
Total	7	14	23	16	60

It might be thought that the scale for salaries ought to be entered as 6, 7, 8, and 9 instead of as 1, 2, 3, and 4, but, as has already been shown in the previous example of the listed pairs, the results will be the same in either case so it is better to use the smaller numbers.

The computation of the coefficient of correlation from the data of this example tabulated in its first form and with the service and salary classes stated in full, gives as a result

$$r=\frac{2,400}{\sqrt{560,000\times 55.6}}=\frac{2,400}{\sqrt{31,136,000}}=\frac{2,400}{5,580}=0.43$$

When the same operations are carried through with the same data tabulated in the simplified form with the classes stated as 1, 2, 3, and 4 the result is

$$r = \frac{24}{\sqrt{56 \times 55.6}} = \frac{24}{\sqrt{3,113.6}} = \frac{24}{55.8} = 0.43$$

While the coefficient of correlation is the same in both cases, it must be noted that care must be taken in stating the meaning of the coefficients of regression. This coefficient for the subject series, when derived from the original tabulations, is $\frac{2,400}{560,000} = 0.0043$ which indicates that on the average in this group the individuals tend to advance $\frac{43}{10,000}$ of a year in service for each additional dollar of salary, or 0.43 years in service for each $100 of additional salary.

If the corresponding coefficient of regression is derived from the data of the simplified tabulation it is $\frac{24}{56}$ or 0.43. The meaning of this is clear if we remember that the 56 in the fraction is derived from the subject series which gives data for salaries, which are in this case stated in terms of steps of $100 each. The fraction then indicates that for each $100 of increase in salary these individuals tend to advance 0.43 of a year in service. The other regression is $\frac{24}{55.6} = 0.43$ and indicates that for each additional year of service they tend to advance 0.43 of $100 or $43.

The foregoing paragraphs may be summarized in three simple statements. The first is that in computing the coefficient of correlation from listed series of paired values the first step is to subtract from each item of each series a number that is smaller by one than the smallest number in the series. The second rule is that in finding the coefficient of correlation from a table of distribution the first step is to substitute for the two series of class values, the numbers 1, 2, 3, 4, etc., being careful to omit any number for which the corresponding class is missing in the original series. The third rule is that in the interpretation of the coefficients of regression it must be remembered that the values are now stated in terms of the class values or steps of the original tabulation.

Editorials

DR. AYRES' FORMULA

A number of inquiries have been made as to the proof of the short method of computing correlation coefficients which Dr. Ayres described in our March issue. The method is entirely rigorous mathematically. Moreover it relates closely to methods already known and used.

Those who work with correlation tables are accustomed to assume averages for the subject and relative series. The deviations are worked out from these assumed averages, and the corrections—i.e., the differences between the assumed and the true averages—are applied. What Dr. Ayres has done is, in effect, to assume an average at zero for both the subject and relative series. The correction or difference between the assumed average and the true average then becomes the true average itself.

Notice how this works out. The formula for the computation of the Pearson correlation coefficient when it is computed from an assumed average may be written as:

$$r = \frac{\Sigma x y - n c_x c_y}{\sqrt{(\Sigma x^2 - n c_x^2)\ (\Sigma y^2 - n c_y^2)}} \quad \ldots\ldots\ldots (1)$$

In this formula c_x is the correction to the assumed average for the subject series and c_y is the similar correction for the relative series. In other words, c_x is the difference between the assumed and the true average for the subject series and c_y is the difference between the assumed and the true average for the relative series. Now when the assumed averages are taken to be zero, c_x becomes the average of the subject series and c_y becomes the average of the relative series. Moreover x and y, being deviations of the measures or items from zero, are the measures themselves.

Substitute, therefore, in the above formula for c_x, A_x and for c_y, A_y where A_x and A_y stand for the averages of the subject and relative series respectively. The formula then reads:

$$r = \frac{\Sigma x y - n A_x A_y}{\sqrt{(\Sigma x^2 - n A_x^2)\ (\Sigma y^2 - n A_y^2)}} \quad \ldots\ldots\ldots (2)$$

Consider now the fact that, by the definition of an average, $A_y = \frac{\textit{tot. of relative items.}}{n}$ If this fraction is substituted in the second term of the numerator in (2), the numerator becomes

$$\Sigma xy - (nA_x)\left(\frac{\textit{tot. of relative items}}{n}\right) \quad \ldots\ldots (3)$$

and the n's cancel. Of course, this is very inelegant; but if, in an effort to be clear to the non-mathematical reader, we carry our inelegance still further, we may (remembering that x and y are now the items themselves) write (3) as:

"Sum of products of subject and relative items
minus
av. of subject items × tot. of relative items"

This is the numerator of the formula as Dr. Ayres gives it (p. 219).[1]

Similar substitutions may be made for A_x and A_y in the denominator of (2). Making these substitutions and noting that an average times the number of items equals the total of the items, we obtain the denominator of the formula as given by Dr. Ayres.

Thus his formula is an application of the usual formula to the case where the assumed average or point of reference is taken as zero. A more mathematical proof might begin by expressing each measure as the sum of the average of its series and its deviation from the average. Thus each measure becomes a binomial whose first term is the average of the series and whose second term is the deviation of the measure from the average. In order to investigate the identity of Dr. Ayres' formula with the usual formula for computing the correlation coefficient, it would next be necessary: (1) to form the products of the successive pairs of measures as written in this binomial form; (2) to square the successive binomials constituting the subject series; and (3) to square the successive binomials constituting the relative series.

The first of these operations gives us a polynomial of four terms for each pair of measures. For the first pair it would be $A_xA_y +$

[1] Of course, we could have substituted for A_x. In that case the numerator of Dr. Ayres' formula would become the "sum of the products of the subject and relative items minus av. of *relative* items × tot. of *subject* items." In other words, the deduction may consist of the product of the average of either series and the total of the other.

$A_y d_1 + A_x D_1 + d_1 D_1$ where A_x and A_y have the same meanings as before and d_1 and D_1 are the deviations of the first subject measure and the first relative measure respectively. Summing the polynomials of this type, we obtain $nA_xA_y + A_y\Sigma d + A_x\Sigma D + \Sigma dD$. Observe, however, that since the sum of the deviations from an average must be zero, the second and third terms of the last expression disappear. Observe also that ΣdD is the same as the Σxy of the usual formula. Evidently, therefore, if we deduct nA_xA_y from this expression, we have the Σxy which we recognize in the numerator of the Pearson formula. Since nA_y equals the total of the relative items, nA_xA_y is "Av. of subject items $\times$ tot. of relative items."

Similarly, the second and third of the operations mentioned above give a series of polynomials of three terms. Their summations yield expressions of the following type (using the notation for the subject series) $nA_x^2 + 2A_x\Sigma d + \Sigma d^2$. Here again, since the summation of deviations is zero, the middle term vanishes. Σd^2 is the usual Σx^2. The reduction must, therefore, be nA_x^2, or $A_x(nA_x)$, or "av. $\times$ total of sub. items." Hence the first quantity under the radical in the denominator of the Ayres' formula proves to be the same as the first quantity under the radical in the denominator of the usual Pearsonian formula. The second quantity proves out in exactly the same way.

The Ayres' formula will therefore yield the same results as the usual Pearson formula, and workers are justified in using it for any advantages which it may have.

B. R. B.

Reviews and Abstracts

BROWN, H. A. *A study of ability in Latin in secondary schools.* Oshkosh, Wisconsin: State Normal School, 1919. 170 pp.

This monograph gives the results of an investigation carried on among the high schools of New Hampshire while President Brown was Director of the Research Bureau connected with the Department of Public Instruction of that state. The tests with the discussions and conclusions based upon them constitute an important addition to the material for measuring the results of high-school instruction. The monograph is for free distribution and may be obtained by addressing the author. A separate edition of the tests is published by the Parker Company, Madison, Wisconsin.

The outstanding feature of the report is the disclosures which it makes of the extremely meager knowledge of Latin which the children had obtained in these New Hampshire high schools. The author believes that the schools are equal to the average of secondary schools, and that the condition as found in them is essentially typical.

At the conclusion of a year's study of Latin, only one-half of the pupils were able to translate sentences as hard as: "*Est copia frumenti in agris nostris.*" After four years of instruction only one-half of the pupils were able to translate a sentence as hard as: "*Dixit aedificium in quo Hannibal esset multos exitus habere.*" The Latin sentence test, although the author supposed that it would prove too easy for the third- and fourth-year pupils, proved on the contrary to be too difficult. This was true despite the fact that the sentences represented the easier ones in beginners Latin books. In terms of medians, first-year pupils were scarcely able to translate two out of the twenty sentences, second-year pupils less than four, third-year pupils less than six, and fourth-year pupils a little over six. The product, therefore, of four years of instruction was the ability to translate six out of twenty sentences representing the easier exercises in beginners Latin books. Since the pupils were given practically all the time they needed, this may be understood to be their best effort.

The investigation might have stopped here. Corroborative evidence, however, was obtained by administering a "connected Latin test," a grammar test, and a vocabulary test. Good scores were obtained only in the vocabulary test; but this was because of its extremely simple character. The five most difficult words were *ipse*, *munio*, *quisquam*, *copia*, and *pax*.

Apparently, therefore, each of the tests tells the same depressing story. It is idle to indulge in arguments as to the "value of Latin" when the evidence is altogether to the effect that children do not learn Latin. The battle of the classicists and the scientists or realists or whoever they are who attack the study of the ancient tongues, suddenly becomes a battle of kites and crows. The issue seems to be of no consequence and the victory not worth fighting for. If Mr. Brown's results are typical, only a very few boys and girls are getting enough Latin to do them any good no matter what its disciplinary, cultural, or linguistic values may be.

B. R. B.

Sixth conference on educational measurements. (Bulletin of the Extension Division, Indiana University, v. 5, no. 1, September, 1919.) Bloomington, Indiana: Extension Division, Indiana University, 1919.

The bulletin, giving the papers submitted at the Sixth Conference on Educational Measurements held by the School of Education of Indiana University, measures up to the quality of its predecessors. Monroe, at that time director of the Bureau of Cooperative Research, states the keynote of the meeting to have been "a plea for diagnosis and more complete interpretation of educational measurements in order that the use of educational tests may result in the greatest possible improvement of instruction." Of the thirteen papers in the bulletin, three by Charters relate to the diagnosis of language and grammatical errors and to scientific curriculum construction, seven by Dr. and Mrs. Pressey and others relate to tests of intelligence, two by Ashbaugh relate to educational service in Iowa and to recent developments in spelling, and one by Monroe discusses next steps in educational measurements.

Charters points out the present composition scales are unanalyzed and hence not particularly useful for diagnostic purposes. He describes in his first two papers his diagnostic language and grammar tests for pronouns, their construction and uses, and the corrective measures to be employed after diagnosis.

The "Cross Out Tests" and the "Primer Scale" for grades I and II are described by Dr. and Mrs. Pressey and their advantages in economizing time and avoiding difficulties in scoring are pointed out. It is interesting to note that this abbreviated team of tests gives a correlation with teacher's rankings of +0.58.

Superintendent Montgomery gives the results of applying Pressey's Mental Survey Test No. 1 at Bedford, Indiana; and State High School Inspector Williams reports the plans for administering the same test to all senior students in the six hundred commissioned high schools of the state. Next to the tests in the army this is the most comprehensive attempt at intelligence measurement recorded and its results will be watched with interest.

Chester S. Carney, consultant with a firm of industrial engineers, at Boston, Massachusetts, reports experiments with mental tests as an aid in the selection and placement of clerical workers in a large factory. A variety of intelligence tests were first tried out on time clerks, but yielded unsatisfactory results. A job analysis of the time clerks' tasks was then resorted to and tests modified or constructed which would probably be diagnostic of ability in these tasks. Mr. Carney concludes that highly specialized tasks may require particular ability which is discoverable only by tests designed for that purpose; and that, while for such purposes general intelligence tests are of little value, they are increasingly useful with reference to the more responsible positions. While personality is not measurable by tests, but must be judged by an interview, "the employer who has once been so fortunate as to obtain and use a really scientific set of tests, will never again be satisfied to pick his men by the haphazard impressions of an interview alone."

Charters, in discussing scientific curriculum construction, attributes the dissatisfaction with the curriculums of the school as not fitting the needs of modern life to three factors: first, the assumption that the best method of equipping for life is to acquaint young people with the organized fields of knowledge according to the systematic arrangement of each; second, the failure to analyze the aims of education until objectives sufficiently detailed have been determined; and third, the failure of educa-

tional psychology to assist the trial and error method of testing by providing a body of facts according to which the material of the curriculum may be arranged in proper pedagogical sequence. For the first method of constructing a curriculum by the impartation of the essential content of each of the highly organized fields of human achievement, Charters argues for a second, "that of studying the daily lives of men to find the most important problems and the selection, from any source available, of such items as will solve these problems." The aims of education are stated in such vague and general terms as to be practically of no use in selecting subject matter. They must be stated in terms of definite working units of achievement. Educational psychology has as yet not given such definite information concerning the interests, abilities, and needs at different ages as to make possible concretely the arrangement of a curriculum in accordance with the psychological nature of children. Various studies, crude and partial, in scientific curriculum making have been made, based on errors or difficulties encountered, on cumulative opinion, on the analysis of projects and on analysis of major objectives.

Monroe, after a brief review of the progress in educational measurements, discusses the next steps under three heads: (1) a critical evaluation of existing educational tests; (2) a more refined and complete interpretation of results; and (3) the use of educational measurements as a basis for modifying instruction.

Ashbaugh in his first paper reviews the five-year history of measurements in Iowa, and in the second the recent work in spelling. This has centered on three problems: (1) the selection of spelling lists; (2) the method of teaching and learning; and (3) the development of scales and tests. The selection of words on the basis of frequency of use in adult correspondence is regarded as the best of the methods so far used. Experimental results are given to show that the method of teaching is not the important thing in spelling, but the method of learning. In measuring results tests and scales are of value for mass results, but individual words and individual pupils may be expected to vary widely from a standard.

V. A. C. HENMON

WHITE, WILLIAM A. *The mental hygiene of childhood.* (Mind and health series.) Boston: Little, Brown and Company, 1919. 193 pp.

This little book does not aim to present an exhaustive analysis of the mental life of the child. It is rather an attempt to put in popular form such principles of the Freudian psychology as may be of value in the training of children. Two conclusions seem to the author to be of especial importance in this connection, one concerning the child, the other concerning the family. "The conclusion concerning the child is that, contrary to generally held opinions, it is possessed of a developing sexuality the roots of which reach back into its infancy. The conclusion concerning the family is also contrary to the opinions regarding that institution commonly held, namely, that there reside within its organization and as a part of its nature certain disruptive tendencies."

The explanation of all conduct is desire growing out of the various forms taken by the two fundamental instincts, self-preservation and sex. The former of these leads to constant attempts to control the environment, physical and social. Those who feel themselves inferior in the struggle for dominance will normally overcome the resulting feeling of apprehension by special endeavor in some field and may become actually superior. The sex instinct frustrated by social conditions of its normal outlet may lead to all sorts of dangerous substitute reactions. On the other hand, if properly "sublimated" they may become the motive for the highest type of human endeavor, such as art and other creative tendencies.

Such a simple statement as the foregoing fails to describe adequately the elaborate mechanisms which according to the psycho-analysts develop in the child's life, and serve to explain his behavior. Some notion of these may be obtained from the chapter entitled "the family situation." The love of the child for the parent has a sex preference—the boy preferring the mother and the girl preferring the father. The parents' preferences are in agreement with those of the child. Under such circumstances the boy becomes jealous of the father because of the attention the latter receives from the mother, and for a similar reason the girl becomes jealous of the mother. At the same time there develops a desire on the part of the child to emulate the parent of the same sex. These emotional relationships to the parents are very important, the child often developing character traits directly dependent upon them.

This description of family life will seem fantastic to those whose experiences have been limited mainly to the observation of normal healthy children, but it is only illustrative of many similar explanations which the Freudians make of human conduct, with sex always lurking in the background. Indeed, one trembles lest the revelation which they aim to make of our inner selves may by the power of suggestion make the facts and the theory somewhat coincide. Of course the Freudian has a ready reply to this implied criticism but that reply has long since failed to convince or frighten.

In at least one respect, however, a book of this sort may have good results. If it serves in any degree to correct our attitude of prudery towards the problems of sex-education, and leads us to face "the original nature of man" with frankness and sincerity instead of hypocritically ignoring it, a useful purpose will be served.

E. H. Cameron

Walsh, James J., M.D. *Health through will power.* Boston: Little, Brown and Company, 1919. 288 pp.

This book presents a striking contrast to the one just reviewed. In the psychology of the psycho-analysts, which the author expressly rejects, a wonderfully intricate and complex group of subconscious processes is invoked to explain mental health and disease. Not so here. Will power explains it all. Does a person suffering from incurable disease live longer than the physician expects or is a person with a frail body able, nevertheless, to accomplish much? The explanation is will power.

Will power has preventive and curative effects on coughs and colds, intestinal disorders, rheumatism, tuberculosis, asthma, abnormal functioning of the heart as well as psycho-neuroses and distinctively feminine ills. One must have the will to eat, to neglect pain, to repress dreads and self-pity, and to take exercise.

Throughout the book the will appears as an occult entity not further analyzable but capable of being injected into almost any situation with marvellous results. "By far the greater part of what has come unfortunately to be called the unconscious and the subconscious and that has occupied so much of the attention of modern writers on psychological subjects is really the will at work." It is the old time faculty psychology in its purest form.

Whatever the objections may be to the book from the more strictly psychological point of view it is but fair to say that it is well worth reading. It is full of a commendable common sense unfortunately too uncommon with regard to health and disease. The chapters on dreads, sympathy, self-pity, and avoidance of conscious use of the will are especially illuminating. The book is written in a style which holds the interest of the reader and should win wide appreciation

E. H. Cameron

SEASHORE, CARL EMIL. *The psychology of musical talent.* Boston: Silver, Burdette and Co., 1919. 288 pp.

This volume contains a message for two classes of persons, the student of the psychology of music and the teacher or supervisor of music in public schools. To the former the book will give a summary of the work of the author and of his students and associates in the psychological laboratory of the State University of Iowa—a laboratory which is "devoted largely to the study of the psychology of music." To the latter it will give scientific information concerning musical talent that has definite practical values.

Professor Seashore has analyzed the capacity required for learning to produce music. This analysis he calls "an inventory of the musical mind." Two coordinate bases of classification are used: first, the attributes of sound (pitch, intensity, duration, and extensity) and second, those traits of the human mind which are necessary for the apprehension and expression of the recognized attributes of sound (hearing of tones, production of tones, the representation of tones in memory and imagination, musical thought, and musical feeling). As a result of this analysis 25 "musical talents" are listed.

As one result of this study of the psychology of music, the author has made available instruments for the group measurement of six musical talents: (1) sense of pitch; (2) sense of intensity; (3) sense of time; (4) sense of consonance; (5) musical memory; and (6) musical imagery. Each of the tests may be given to a group of pupils in a half-hour period. The author recommends their use in the fifth grade as a preliminary survey and again in the eighth grade in order to advise pupils concerning the avocational and vocational pursuit of music.

Possession of these talents is a prerequisite for certain types of musical achievement. Since, for the most part, they represent native equipment rather than the results of training, measures of them in connection with other information furnish a very reliable basis for the vocational and avocational guidance of school children. The furnishing of a basis for such guidance rather than the measurement of musical ability or of the results of teaching is the real function of the tests.

The volume is a very important contribution in the field of applied psychology. Concisely stated paragraph headings materially assist the reader. Portions of the book, however, do not appear to be well balanced because of the brevity of the treatment of certain topics and the elaboration of others. This is no doubt due in part to the undeveloped character of the field.

In general the book is weak in the presentation of data, but strong in keenness of analysis and interpretation. The graphical charts are below present-day standards. Practically none of them can be read without looking up the notation in the text, and in many cases it is necessary to turn to another chapter to get this information.

It is seldom that a fundamental treatise, such as this volume is, fails to include a carefully prepared bibliography. This book contains no bibliography and the footnote references are frequently of the "*op. cit.*" type. Neither is there an index. The elaborate topical table of contents partially takes the place of this but not entirely. The responsibility for these defects and the one concerning the graphical charts should be shared by the editor and publisher of the book. It is announced as the second volume of the Beverly Educational Series and it is hoped that such defects will be eliminated from future volumes.

W. S. M.

News Items and Communications

This department will contain news items regarding research workers and their activities. It will also serve as a clearing house for more formal communications on similar topics, preferably of not more than five hundred words. These communications will be printed over the signatures of the authors. Address all correspondence concerning this department to Walter S. Monroe, University of Illinois, Urbana, Illinois.

Idaho Bureau at Lewiston

Director Chas. L. Harlan of the Bureau of Educational Tests, Standards, and Measurements writes that he has just returned from a tour through the state of Idaho in which he gave, scored, tabulated, and interpreted both educational and intelligence tests in ten cities and to more than twenty thousand children. The tests used were Monroe's Standardized Reasoning Tests in Arithmetic, Courtis Standard Research Tests, Series B, Charters Diagnostic Tests in Language and Grammar, Woody Arithmetic Scales, Haggerty Intelligence Examinations, and the Otis Group Intelligence Test.

Directors of Supervised Student Teaching

At the recent meeting of the Department of Superintendence at Cleveland a new educational organization was launched that is known as the National Association of Directors of Supervised Student Teaching. The purpose of the organization is expressed as follows in the preliminary announcement.

"The dominant purpose of the association will be to test out in the training school the principles of psychology and education taught in the professional courses offered in the training department of the institution, with the view of eliminating such instruction as will not stand the test of sound practice. This will mean the complete unifying, harmonizing, and simplifying of the whole scheme of supervised student teaching."

Southern California Research Association

The Southern California Educational Research Association was organized in December 1920. Its purpose is stated as follows: "(1) To discuss the value and use of standardized tests and statistical measurements; (2) To construct new tests and measurements as the need arises—these to become the property of the Association and not for private gain; (3) The study of such other phases of Educational Research as may be desired by the Association."

It was decided to hold meetings every two months in Los Angeles, and the following officers for the year 1920 were elected: Dr. W. Franklin Jones, President; Dr. A. H. Sutherland, Vice-President; Miss Mary Bess Henry, Sec'y-Treasurer. A charter membership of forty-three seemed very encouraging for the future of the association.

Mr. Ernest Branson, Director of Research at Long Beach, spoke at the January meeting on "The Organization of a Department of Research in a Small City." Mr. Branson also gave an interesting report on the results of intelligence and educational tests in the Long Beach schools.

March the twentieth, Dr. J. Harold Williams, Director of Research in the Whittier State School, gave an excellent talk on "Clinical Psychology in the Public Schools."

The closing meeting of the year will be in charge of Dr. A. H. Sutherland, Psychologist of the Los Angeles City Schools, who will speak on some important adaptations of clinical psychology in the Los Angeles schools.

Batteries of Tests for School Surveys

The systematic survey of schools or school systems by means of standardized tests is a common method of appraising "efficiency." Relatively little has thus far been done, however, by way of putting together in convenient form groups of tests for this type of use. In planning for a survey the superintendent usually orders a variety of tests from a variety of sources; and as a consequence he often gives them in varying order. The directions from test to test are not related, sizes of blanks often vary; and, in order to assemble the results of an extensive survey, several separate blanks must be handled for each child. In order to make the procedure in survey work more systematic and convenient, Dr. and Mrs. Pressey of Indiana University have recently put together, in four-page folder form two batteries of standardized tests: (1) an examination for the junior high school and (2) an examination for the second grade.

1. The examination for the junior high school consists of four tests, one on each page of the folder. The first page gives a test in American history, the questions being selected from the Bell and McCollum test material.[1] The second page consists of the Starch Arithmetic Test A,[2] slightly rearranged for use in grades six to eight. On the third page is presented an examination in geography, the questions being selected from the Hahn-Lackey Geography Scale. The fourth is a test of English.[3] The scale thus gives one test in each of the four "promotion subjects." The timing is very simple, eight minutes to each test.

2. The second-grade scale consists of a similar folder. On the first page are spaces for writing in twenty words from the Ayres spelling list. The second page presents a test in reading vocabulary. It consists of 24 lines, in each of which there is one real word and four nonsense syllables, thus, "suntrel winter bolun omutse ortuni." The child is told to "find the real word and draw a line around it." The third page contains two tests, one in the fundamental addition combinations and one in the corresponding subtraction combinations, the test being exactly similar to tests A and B of the Cleveland Survey Scale.[4] On the last page is a test of the understanding of sentences; there

[1] Bell, J. C. and McCollum, D. F. "A study of the attainments of pupils in United States history," *Journal of Educational Psychology*, 8:257–74, May, 1917.

[2] Starch, Daniel. "Scale for measuring abilities in arithmetic," *Journal of Educational Psychology*, 7:213–22, March, 1916.

[3] The test consists of nine compositions written by school children, selected from Thorndike's list of 150 compositions, published by Teachers' College in 1916. The compositions have all been rated on the Hillegas Scale and have values running, by fives, from 30 to 70. The compositions are presented in a random order and the children are told to number them in order of value. The score is in terms of deviations from the correct order.

[4] Judd, C. H. *Measuring the work of the public schools*, pp. 95–96.

are again 24 lines, each consisting of a sentence with one word too many, thus, "She likes clock to swing up and down." The child is to cross out the extra word.

The two batteries of tests have been carefully standardized by use in the survey of the school systems of three Indiana cities and may now be obtained for general use. The blanks have a decided convenience as there is only one blank to handle for each child; and they make a survey more systematic, since the tests are always given in the same order. And, by a simple transmutation and combination of the scores on the separate tests, an "achievement total," or single numerical statement of general school attainment, may be obtained for each child.

The examinations are designed primarily for school survey work. However, the authors have conceived them as particularly useful for certain other purposes. (1) They should give a means of measuring, not merely progress in a particular subject, but attainment as a whole; and they should thus make possible a synthetic statement of the educational status of each child, in a much larger and more comprehensive way than would otherwise be possible. (2) It should be similarly possible to appraise teaching, as a whole. (3) Finally, by the combined use of such scales and a scale of intelligence it should be possible to set over against each other general ability and total achievement and hence to evaluate what a child has done with relation to what he should be able to do. Similarly it is possible to evaluate a teacher's success with due regard to the quality of "pupil material" with which she has had to work.

The Scores of a Group of Rural Teachers on Courtis Standard Supervisory Test in Geography

Is there a minimum standard which we have a right to set for rural teachers in locational geography of the United States? Should a teacher who expects to teach all the geography from first grade through the eighth be able to locate any state of the Union on the map? Should she be able to tell the state in which the thirty principal cities of the country are located, not by naming the state, but by placing a number opposite the name of the city which number occurs on the state in which the city happens to be?

Without attempting to answer the questions above, I desire to give the results of a test of the performance of a group of about one hundred rural teachers in an Iowa county which I must not call typical. It is not typical in that it contains a teacher training institution, and has within its borders one of the leading industrial cities of the Middle West, as well as another good sized town of several thousand people. Such a county is not, of course, typical of Iowa counties. The Courtis Standard Supervisory Test in Geography was given at a teachers' institute. The teachers were not asked to sign the papers. Eighty-six papers were returned. On the location of the forty-eight states the results were as follows. The median number of states located correctly is 40, the extreme variation being from 13 correct replies to perfect scores, of which there were two. Of the total,

2	graded	between	10	and 14
5	"		15	" 19
10	"		20	" 24
8	"		25	" 29
7	"		30	" 34
11	"		35	" 39
24	"		40	" 44
19	"		45	" 49

Twenty-eight papers were unfinished. With these, all states after the last one marked in the list are not counted "omitted." The District of Columbia is most frequently omitted, having been avoided by 32 writers. It was misnumbered by 18, most of these mistakes being the placing of the District in either Maryland or Virginia, usually the former. (However, one teacher placed Washington in Iowa. And of course, there is a Washington in Iowa.)

Of the states, Delaware is the most puzzling Will o' the wisp. It was omitted by 9 teachers, and missed by 28. It seems to be confused with New Jersey very frequently. This usually makes for a mistake in the placing of New Jersey, which is second in the list, with 30 errors. Another eastern state, New Hampshire, shares third place with Colorado and Kansas, each with 28 errors. New Hampshire is often interchanged with Vermont, on which there were 25 mistakes.

I shall not go into detail as to number of errors with other states. There were 16 missed 20 times or more, while 11 were missed 10 times or less. California, Maine, and Florida were each missed by one teacher. Iowa, strange to say, was misnumbered or omitted by 4. States contiguous to Iowa stand low in the scale of errors, Minnesota with 5, Wisconsin with 8, Illinois with 2, Missouri with 14, Nebraska with 14, and South Dakota with 16.

It is evident that with some states there is ease in remembering "outstanding" locations—"something to fasten to." Witness California, Maine, Florida, each missed once, and Texas not all. Rhode Island was misnumbered 22 times.

On the whole, the errors of teachers seem to fall into two comparatively easily distinguishable classes; first, the interchange of poorly oriented outlines, as Vermont and New Hampshire; Connecticut and Massachusetts; Washington and Oregon; New Mexico and Arizona; New Jersey and Delaware; second "random shots" or "bonehead mistakes"—to use a term from athletics. There is often indication that the subject has no idea as to location, making a mere guess at it; while in other cases he seems to be under a mistaken impression, perhaps judging more by shape than by locality.

Among the cities, Washington is most often misplaced or omitted, having been missed by 49 teachers. Many of these errors are due to placing the capital city in Maryland, which shows a knowledge of its whereabouts, but not an accurate one. Savannah stands second, with 35 errors, and Duluth third with 32. No teacher mislocated or omitted San Francisco; all other cities were misplaced or omitted by at least 3 people. Mislocations of cities sometimes correlate with errors in states, and sometimes not. For instance, one teacher gives Kansas the number properly belonging to Nebraska; she is likely to place opposite to Kansas City the Nebraska number which she mistakes for that of Kansas. Many writers placed Rochester in Minnesota. This was accepted for the great hospital at that point has given the Minnesota Rochester a reputation equal to that of the larger city of New York, and superior to it in the Middle West. Strange to say, some people make mistakes in numbering a state, but when they come to a city within that state, get it correct.

The scores for cities were as follows, in groups of 2:

2	scored	7 or 8
1	"	9 or 10
1	"	11 or 12
5	"	13 or 14

3	"	15 or 16
13	"	17 or 18
4	"	19 or 20
5	"	21 or 22
12	"	23 or 24
18	"	25 or 26
14	"	27 or 28
8	"	29 or 30

The median is 23 or 24.

I expect to continue this work as opportunity offers; and to suggest methods of improvement to teachers.

FRED D. CRAM
County Superintendent, Mason City, Iowa.

Wanted: A History Scale Maker

Dr. Buckingham's article, "Index of Efficiency in Teaching United States History" in the March number of the Journal challenges the expert scale maker.

"May I not" assume that his core-syllogism is something like this, disregarding the preliminary analogies from science:

(a) The Champaign-Danville tests were Van Wagenen's.

(b) Van Wagenen's questions are nearly all essentially fact questions, etc., and, inferentially (p. 169) are not adequate for testing the "higher abilities" they presume to be designed for.

(c) Correlations made show that knowledge of facts is symptomatic of "higher abilities." (Knowledge of facts "conditions" these higher abilities. Granted. Unanimous vote.)

Therefore: Fact question scores correlate positively with pseudo-think-but-really-fact question scores.

(The examination of scores made in New York and Madison was the product of test-questions that Dr. Buckingham did not describe, and therefore cannot fairly be considered here.)

The conclusion is penultimately stated on p. 169—the "even-bet" phrase tempts me to say "ante-penultimately." It is: "Whatever may be the desirable outcomes of the teaching of history, it does not appear to be necessary that a valid test should require reactions with reference to all of them. It may merely require reactions which are symptomatic of them."

Does this not challenge discussion if not dissent? In the first place (the "even-bet" phrase again lures) is it not a case of passing the buck? Is not the inference just this: that "Van Wagenen is the first maker of a thoroughly worked out scheme" etc.; that the use of his largely pseudo-think test questions *indicates* that correct fact answers are symptoms of higher abilities, *but not valid evidence* of the same; and that so far we have no series of valid tests for the purpose.

The writer, of the laity so far as the N. A. D. E. Research is concerned, persists in his innocent faith that here is a field for work—that valid tests *can* be made; that at least, questions so far removed from the abstract factual as to justify an examination of the relation between memory of facts and *actual* attainment of the higher abilities, can be framed.

If I am correct in this, and there *can* be shown a fairly high positive correlation, *then* we shall indeed "be justified in measuring historical ability by means of the ability to answer fact questions" (p. 165). And if there *cannot* be shown a high positive correlation, then two series for the two different main objectives—facts and higher abilities—should be used, or at least two sets of measures obtained from one series of questions.

Now this presupposes as Dr. Buckingham implies, that we must define "historical abilities" and agree upon the definition. Without now suggesting a definition, let us see if one cannot be "felt" or sensed in the examination of a few questions and topics—for the writer, not of the laity so far as teaching history is concerned, has preponderant proof from his own experience and that of his teachers that "if, after giving a fact test, we should desire to know how capable the pupils probably would be in other (i.e., higher abilities) ways," we may *not* infer this "without giving any other test." Fact knowledge does not always or necessarily "carry over" into historic perspective, sense of significance of events, etc., etc. (p. 167).

Consider the following question: (A) Daniel Webster was at one time against a protective tariff and later favored protection. John C. Calhoun at one time was a protectionist and later vehemently opposed protection. Jefferson was a strict constructionist and yet favored the purchase of Louisiana; the Federalists were liberal constructionists, and yet opposed it. The Democratic party until recently was a states rights party; and New England republicans until recently were opposed to the states rights idea. What conclusion do you draw regarding political principles and theories?

Here is a question in which the *facts* are furnished gratuitously. The "higher ability" is to induce the law that—politics is full of inconsistencies; that political theories are really almost solely actuated by personal, local, selfish interests.

(B) Or take this: Suppose Champlain in 1608 had chanced to befriend the Mohawks (Iroquois). What would probably have been the results of the New York campaign of the French in 1758?

(C) Or take any text—Woodburn and Moran, Forman, Hart, Fite, any of them—and examine the "suggestive questions" usually given. Can it not fairly be said of many of them that (1) some are 90 percent fact; (2) some call for "thinking which has no *necessary* reference to history"; while (3) others demand thinking which does make a demand upon the "higher abilities"?

The answers may be *fact* answers, to be sure. But the pupil must go-a-fishing for them; he must examine, weigh and accept or reject facts; he must then organize them. And that requires staunch thinking.

Do these facts that he selects have "a *necessary* reference to history?" It would appear so.

From which I *persist* in my innocent faith that a job is looking for a taker; and *insist* that it would be very unsatisfactory to test for facts only and continue to *estimate* subjectively by inference of problematic symptoms, the capabilities of pupils to achieve the "higher abilities" that good history teachers desire for them.

F. S. CAMP
Superintendent of Schools, Stamford, Conn.

National Association of Directors of Educational Research

(E. J. Ashbaugh, *Secretary and Editor*)

The last issue before summer vacation. The past year has certainly been one of great advance with the members of the association. Every letter from individual members, as well as conversation with those whom the secretary has seen, has breathed the spirit of optimism regarding the progress of our work. The fields of activity open to us, whether working in a city system or in a state-wide, are a challenge to the utmost effort which we can put forth. Most of us will have closed the regular year's work before the coming of this issue of our Journal. Let us summarize the results of the past year, lay out a big program for the coming year, and notify the secretary that these programs may be printed in the first issue of the Journal next fall.

Only a few members have sent in any statement of their recent work. Miss Brown, Director of Measurements in Denver, reports particularly interesting work in intelligence testing in the primary grades and in the junior high school. She has also given the Henmon Latin Test to all pupils who have had one or more semesters of Latin for the purpose of making a comparison of the teaching of first-year Latin in junior and senior high schools. City-wide tests in reading and arithmetic given in the first semester are to be repeated in May.

Miss Mallory at Beaumont, Texas, reports the repetition of standard tests with about twenty-five hundred children. This is the first testing that has been done in Beaumont and some very interesting results have been secured. Both Miss Mallory and Superintendent Moore feel that the tests given during the first semester gave them a better insight into the condition of the schools, than they could have secured in the same length of time in any other way.

Professor W. A. McCall reports that he will have an article in the next issue of the *Teachers College Record*, entitled "How to Compute the Median." This will doubtless be of great interest to members of the Association, especially in view of the fact that a report was made by Mr. Courtis at the Cleveland meeting in behalf of the committee on standards which took up this same problem.

Dr. Woody of the University of Washington reports that he has received material for various phases of a reading investigation carried on in six of the largest cities of the first class in the state of Washington. Monroe Silent Reading was given to all the children in the systems and Thorndike Paragraph Reading to children in three buildings in each city. An extensive questionnaire on school room practice was sent out at the time the testing was done. In addition to this he is still continuing his investigation of reading-arithmetic exercises and is attempting to develop some school tests for differentiation of pupils in the seventh grade, according to ability.

Dr. Greene and your secretary have conducted an extensive investigation of the salaries of Iowa teachers in relation to their preparation, experience and tenure. A report on salaries was sent to each member of the association some time ago. A report on the relations of salaries, training, experience and tenure was made at the Sixth Annual Conference on Supervision held by the University of Iowa this spring. The report will be published in the organ of the State Teachers Association and a copy will be sent to each member as soon as it is in print.

We are also collecting the results of testing over the state with the four most widely used standard tests. These tests are: Courtis Series B Arithmetic, Monroe Standardized Silent Reading, Courtis Silent Reading, and Greene's Organization. Tabulations for the state on these tests will be available next fall.

An extensive study of teachers' salaries in Wisconsin has been made by Dr. Theisen and Mrs. Flemming. They have doubtless been busy on many other problems, but have failed to report to the secretary in detail.

If the association has failed in any sense this year, it has been in the fact that we have not kept closely enough in touch with each other. From a few members repeated calls from the secretary have failed to secure a response regarding their work. Of course you are busy. But do you not like to know what the other members are doing; what tests they are trying out, and what unique things they are attempting to do? Of course you do. So would all the rest of us like to know more about what you are doing. Let each one resolve to furnish the secretary, not only with the program asked for above, but also with one or two reports, at least, during the year of progress on these programs.

VALUABLE ARTICLES TO APPEAR IN EARLY ISSUES OF THE JOURNAL OF EDUCATIONAL RESEARCH

Comparative Results in Intermediate and Elementary Schools at Los Angeles by *Thomas H. Briggs.*

Survey of the Reading Tastes of Children by *Harriet S. Wardell.*

Grade Norms for the New York City Penmanship Scale by *Eugene A. Nifenecker.*

Use of Group Mental Tests in the Guidance of Eighth-Grade and High-School Pupils by *Virgil E. Dickson.*

Provisions for Individual Differences in the Teaching of Reading by *W. W. Theisen.*

Use of Psychological Tests in Vocational Guidance by *W. M. Proctor.*

Standardizing Procedure in Combination Systems of Supervised Study by *W. H. Hughes.*

Geography Testing in Boston by *Harriet M. Barthelmess.*

Third article on "Putting Standardized Tests and Scales to Practical Use in Rural Schools" by *S. S. Brooks.*

Scale of Attainment No. 1 by *Luella W. Pressey.*

Educational Measurement in the Field of History by *Charles Harlan.*

Comparative Study by Educational Measurements of One-Room Rural-School Children and City School Children by *J. C. Chapman* and *E. L. Eby.*